SPATIAL VARIABILITY OF SOIL VELOCITY USING PASSIVE SURFACE WAVE TESTING

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Civil and Environmental Engineering

by

Daniel Raymond Wagstaffe

December 2015

COMMITTEE MEMBERSHIP

TITLE:	Spatial Variability of Soil Velocity using Passive Surface Wave Testing
AUTHOR:	Daniel Raymond Wagstaffe
DATE SUBMITTED:	December 2015
COMMITTEE CHAIR:	Robb Moss, Ph.D., PE Associate Professor of Civil Engineering
COMMITTEE MEMBER:	John Jasbinsek, Ph.D. Associate Professor of Physics
COMMITTEE MEMBER:	Gregg Fiegel, Ph.D., PE, GE Professor of Civil Engineering
COMMITTEE MEMBER:	Dmitry Maleev, Ph.D. Candidate of Geology and Mineralogy Associate Professor at Far Eastern State Transportation University

ABSTRACT

Spatial Variability of Soil Velocity using Passive Surface Wave Testing

Daniel Raymond Wagstaffe

Lifelines such as highways, pipelines, telecommunication lines, and powerlines provide communities with vital services, and their functionality is dependent upon the soil that supports them. However, when designing the infrastructure, it can be difficult to know where to test the soil in order to give spatially representative sampling, particularly for long, lifeline structures. Finding this distance requires knowledge of the spatial correlation and/or the spatial variability of the soil parameter (stiffness, cohesion, etc.). But this correlation distance is not typically found in practice because it requires large amounts of data and the costs of retrieving that data can be high. Lack of representative sampling can lead to an overly conservative design but too much sampling can create an overly expensive sampling program. In this study, multiple tests using the geophysical method of spatial autocorrelation (SPAC) were conducted to find soil velocity along a 310 meter long profile. SPAC records passive surface waves which sample the underlying soil, and these surface waves can be used to create a shear wave velocity profile of the site. The spatial continuity of the stiffness (the soil velocity values) was then found using geostatistics. The geostastical tool primarily used in this study was the (semi-)variogram, but the covariance function and the correlogram are also shown. The distance that the soil parameter is minimally correlated with itself is assumed to be the maximum distance that gives representative sampling. This study found this distance (the range of the semi-variogram) to be 70 meters for 5 meters depth, 100 meters for 10 to 15 meters depth, and 90 meters for 30 meters depth.

ACKNOWLEDGMENTS

First and foremost, I would like to thank Dr. Robb Moss. This study would not have been possible without him and I very much appreciate all the help he gave me. I always got undivided attention when I came in for advice and I never got shooed out the door. I am grateful that he not only helped develop this study, but also helped grow me as a student and an engineer.

My other committee members who each played a role in the development of this thesis. To Dr. John Jasbinsek who was eager to send reference material, enthusiastic to sit with me and discuss my newest questions, and willing even to put on gloves and help me collect the data. To Dr. Dmitry Maleev who was ready to give advice and guidance even before his office was set up and for the much needed help in translating the Russian code book. And to Dr. Gregg Fiegel who developed my engineering skills in previous classes, and who took the time to read through and constructively critic this thesis.

To my friends and roommates for working alongside me, for encouraging me, and for always providing the smile I needed (and the food I needed after long nights in the lab). To Tyler, Jordan, Luke, Rory, Nick, David, Riley, and Ballard: thank you.

And, of course, to my family. For pushing me and loving me since I was too young to know what engineering was and when I thought that a thesis defense had something to do with a dragon and a sword. I am the man I am today because of your love and guidance. Thank you.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1.1 Overview

Lifelines do not typically have the benefit of being confined to a small area. Their long, linear nature implies that they will cover more ground, and this creates more opportunity for changing soil conditions beneath them. When designing these structures, it is important to know where to group similar or correlated material. The similarities and dissimilarities will govern where to test/sample the material and possibly where to change the design.

Europe and Russia both have codes that give guidance on these correlation distances. The European code (EN 1997-2: Eurocode 7: Geotechnical Design – Part 2: Ground Investigation and Testing, 2007) gives broad guidelines. In chapter 3, they recommend sampling based on investigations, the geology, and the complexity of the structure, but they do provide recommendations in Annex B.3 based on the type of structure. For linear structures such as "roads, railways, channels, pipelines, dikes, tunnels, [and] retaining walls," they recommend a sample spacing of between 20 and 200 meters. The Russian code (Engineering Surveys for Construction. Basic Provisions, 1996) has more specific guidelines for linear structures. They recommend the bandwidth zone (the transverse distance that points can be sampled between), the longitudinal distance between sample locations, and the required sample depth. A summary of the recommendations from the Russian code can be seen in Table 1.1. Note that this study is primarily interested in the longitudinal distance between samples, and this table shows

the range to be between 100 and 1,000 meters (or 100 and 500 meters if power lines are ignored).

Type of Linear Object	Bandwidth of Investigation Zone (m)	Longitudinal Distance between Investigation Points (m)	Sample Depth (m)	
Railway	200-500	350-500	Up to 5	2 m lower than the standard depth of soil freezing
Highway	200-500	350-500	Up to 3	
Main Pipeline	100-500	300-500	At 1-2 m below the depth of the pipeline	
Overpass for Terrestrial Communication	100	100-200	3-7	
Power Lines (kV):				
Less than or equal to 35	100-300	500-1000	3-5	
More than 35	100-300	500-1000	7-10	
Cable Link	50-100	500-1000	At 1-2 m below the depth of the pipeline	At 1-2 m below the standard depth of soil freezing
Water Supply, Sewerage, Heating and Gas Pipelines	100-200	100-300		
Underground Collectors - Gutter and Communication	100-200	100-200	2 m below the estimated depth of the reservoir	

Notes

1. Minimum distances should be used for complex geotechnical conditions and the maximum distances for simple geotechnical conditions.
2. In areas with specific soils, the development of dangerous geological processes should reduce the distance by three to five workings.
3. If the corridor is supposed to trace the design of several linear features, the number and depth of excavation is set based on the minimum distance and maximum depth for the corresponding linear objects.

Table 1.1 Typical design for bandwidth, sample spacing, and sample depth according to Russian code

America has yet to come up with recommendations for this sample spacing. This study's goal is to give a recommendation for this spacing. To do this, we collected velocity values in alluvial soil and found the correlation distance. The data was collected using the spatial autocorrelation method (SPAC) which gives Rayleigh wave and shear wave velocity values. These velocity values were then analyzed using spatial statistics to

find the correlation distance. Since the correlation distance denotes when the soil becomes dissimilar, this distance can be assumed to be the maximum, allowable sample spacing.

1.2 Organization of Thesis

This thesis is written with the hope that any person with a moderate, technical background can understand the work. Background material is presented before delving into the results in order to provide the reader with an understanding and context of both surface-wave testing and geostatistical tools. The first part of chapter 2 reviews the concepts behind surface-wave testing, how to collect and analyze this kind of data, and some limitations behind the methods. The second part of chapter 2 reviews the concepts of geostatistics, its common functions and models, and some limitations associated with these tools. Chapter 3 goes through the hands-on aspect of surface-wave testing. It gives instructions on what equipment was used, how to use the equipment in the field, and step-by-step instructions on how to process the data. Then the results are presented in chapter 4 followed by the conclusion and recommendations for further research in chapter 5.

1.3 Geology of the Site

An overview of the site's location (Cuesta College's campus) can be seen in Figure 1.1 (Google Maps, 2015). The surficial geology of this site and the surrounding area can be seen in Figure 1.2. It shows that the project lies on young, alluvial flood-plain deposits (Q_{ya}) whose description is also shown in Figure 1.2. In the surrounding area, all of the Cretaceous to Jurassic units (KJ) shown in the figure are of the Franciscan complex. They consist of mélange (KJ_{fm}), sandstone and shale (KJ_{fss}), metavolcanic rocks (KJ_{fmv}), and chert (KJ_{fch}). Other nearby units consist of serpentinized, ultramafic rocks

(J_{os}), a volcanic intrusive complex (T_d), landslide deposits (Q_{ls}), alluvial flood-plain deposits (Q_a), and felsite (T_f).

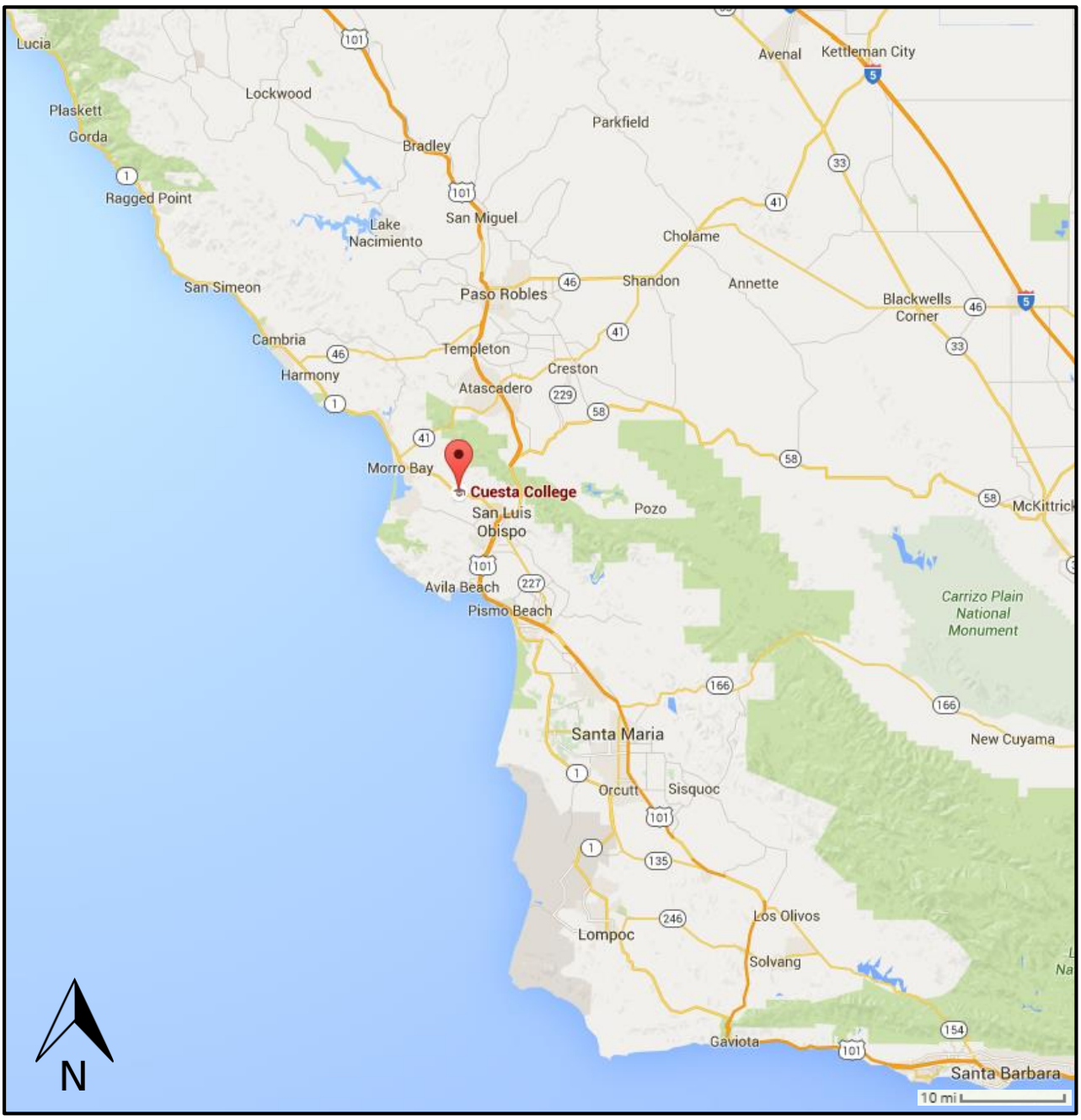


Figure 1.1 Vicinity map of the site

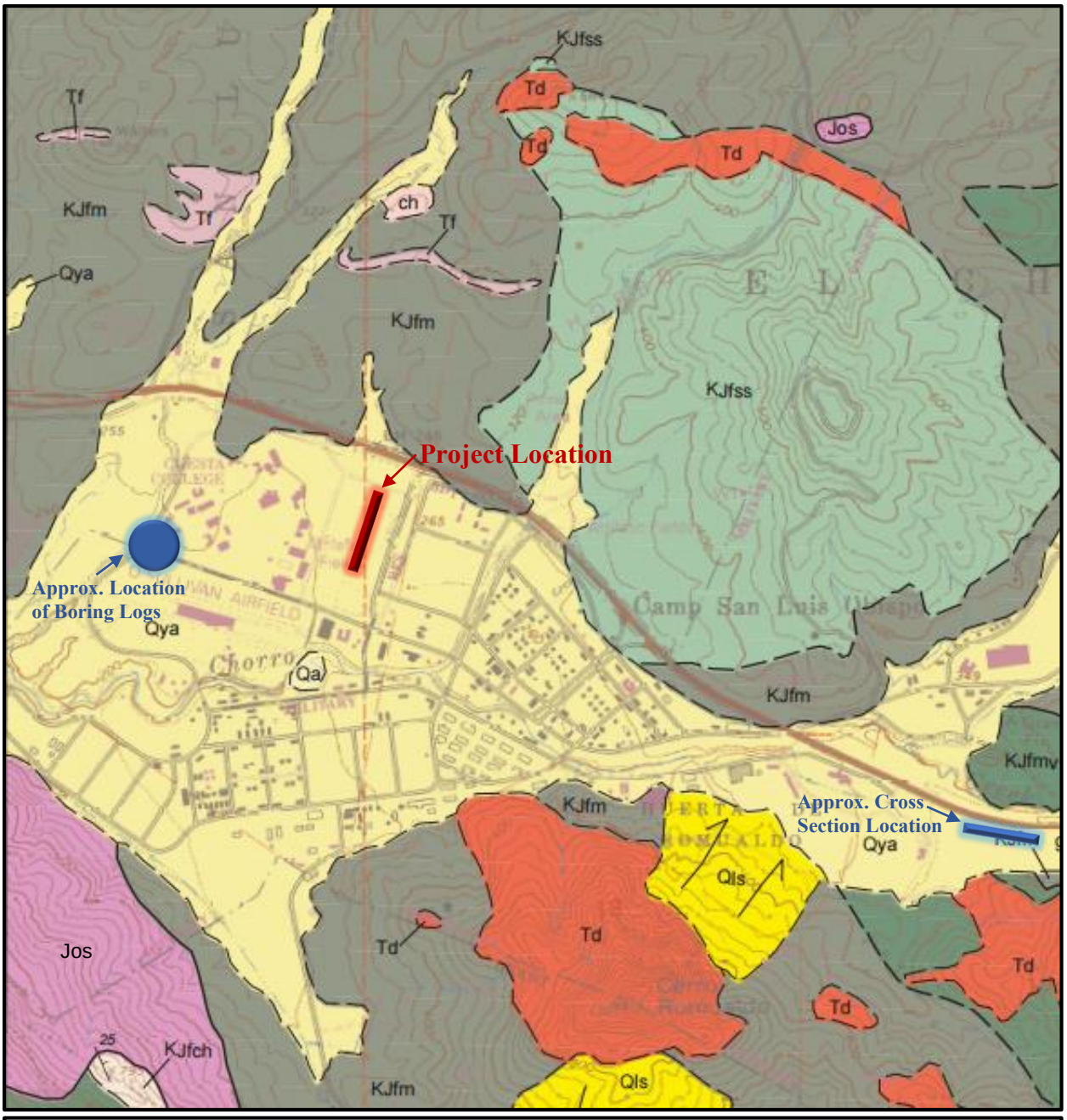


Figure 1.2 Geologic map of the site and the surrounding area (California Conservation and California Geologic Survey, 2010)

Chipping (1987) describes these units in detail. He writes that the alluvium is no thicker than 80 feet near the coast and thins to "probably" no more than 20 to 30 feet as one goes further inland. These sediments are relatively fine. He also explains that the Franciscan formation mélange is the most common rock form in the area which consists of a mixture of sandstone, chert, serpentine, basalt, greenstone, shale, and high-grade metamorphic rocks. The volcanic rocks are nearly all basaltic and can be found as blocks within the mélange or slab-like masses in the mélange that can stretch for miles.

Earth Systems Pacific also provided a nearby, geologic cross-section and eight boring logs from a project at Cuesta College. The cross section (from a report submitted in 2015) can be seen in Figure 1.3, and its approximate location can be seen in Figure 1.2. The boring logs are from a 2010 report and can be seen in the Appendix. Their approximate location can also be seen in Figure 1.2. The cross section shows alluvial material with a thickness ranging from approximately 30 to 60 feet (9 to 19 meters) being underlain by metavolcanic, serpentinite, and sandstone bedrock. The boring logs are all drilled to 11.5 feet (3.5 meters) and show sandy clay with some locations underlain by claystone (Franciscan formation). The three logs that show claystone give depths of 4 to 7 feet to this layer. The other five logs show sandy clay for the full 11.5 feet.

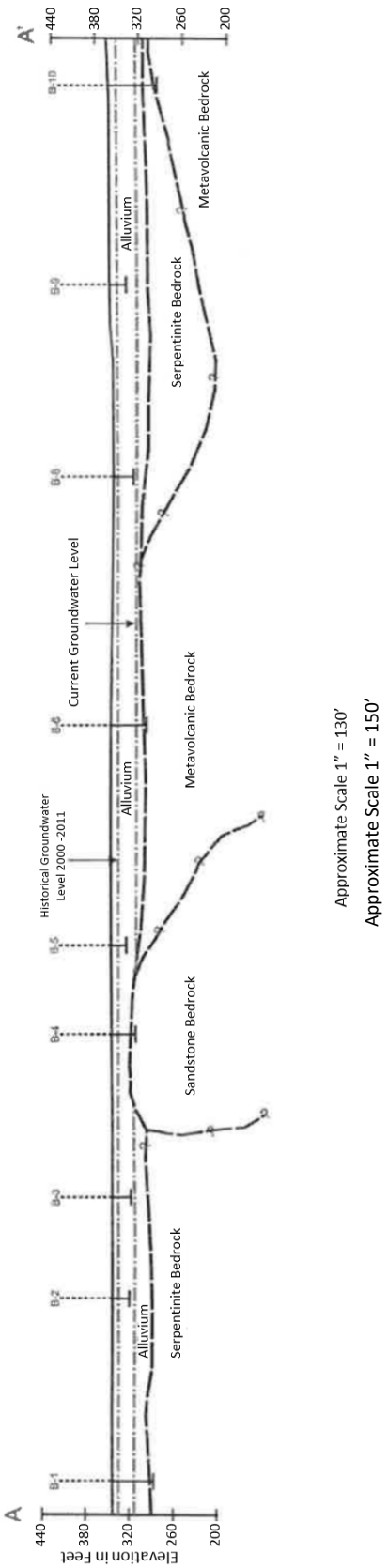


Figure 1.3 Cross section of nearby location (provided by Earth Systems Pacific)

Chapter 2: Literature Review

The aim of this chapter is to give background knowledge of both the spatial autocorrelation (SPAC) method and of geostatistical tools so that a framework can be established before delving into the study's results. This overview is inherently broad due to the complexities of the subject matter, but references to other works are included if further depth is desired.

2.1 Review of the Spatial Autocorrelation (SPAC) Method

2.1.1 Introduction to SPAC

There are two types of seismic wave testing – invasive and noninvasive. Invasive testing acquires data from within the geologic material whereas noninvasive testing relies on data retrieved from the ground surface. Cross-hole, down-hole, sCPT, and suspension logging are examples of invasive testing. Noninvasive examples include refraction, reflection, and surface-wave testing. The benefit of invasive testing is that the uncertainty of the values (usually seismic velocity) is smaller since the data is recorded from inside the material, but it can be expensive to explore large areas. The opposite is true for noninvasive methods. At a cheaper cost, they can sample large amounts of material, but the values have greater uncertainty (Moss, 2008).

SPAC is a noninvasive test that was first introduced by Aki (1957, 1965). Aki took the idea of characterizing the earth's crust with earthquake tremors, and formed a method to characterize the earth's near-surface material by using microtremors. For SPAC, these microtremors consist of passive, ambient noise such as highway traffic, wind currents, and ocean waves. By recording these microtremors with a two-dimesional

array, the direction of the wave propagation can be found and the speed of that wave can be used to estimate properties of the underlying geologic material.

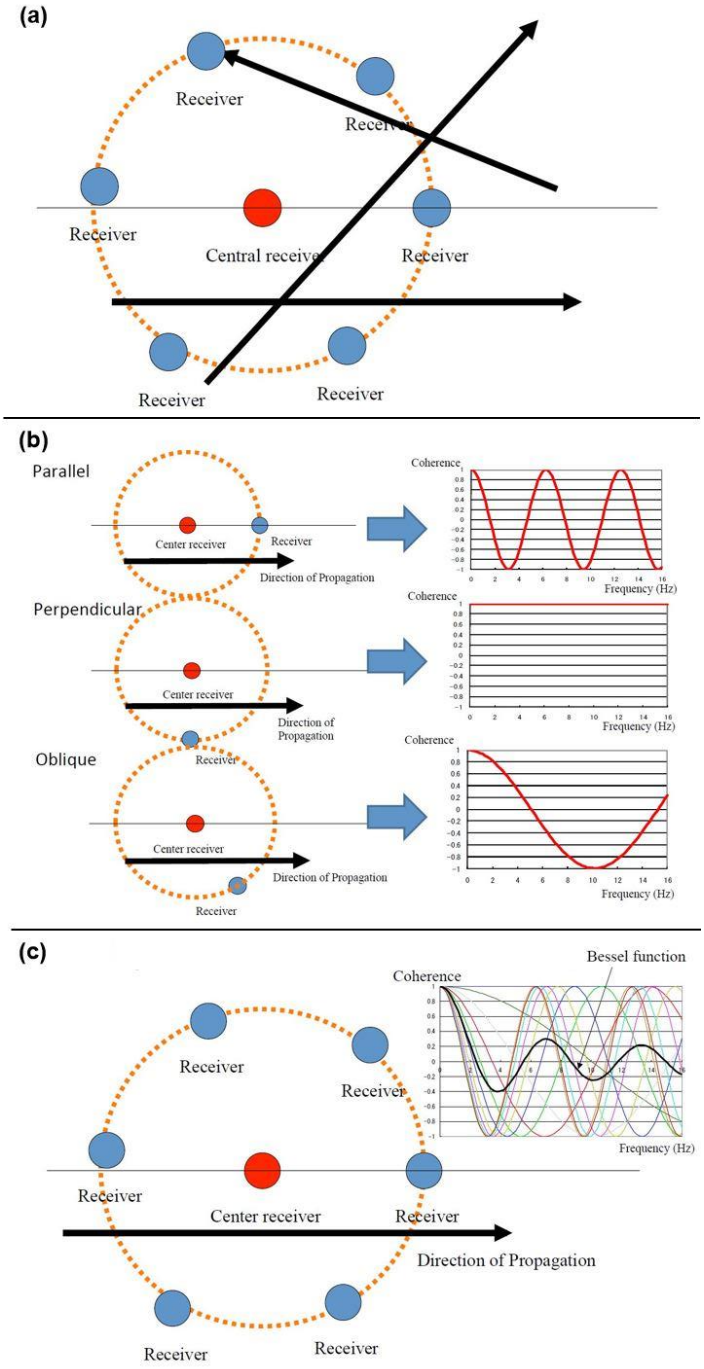


Figure 2.1 The derivation of spatial autocorrelation as presented by Kiochi Hayashi; (a) Waves propagating in different directions, (b) the coherence of the wave propagation between pairs of receivers, and (c) averaging the coherence between all receiver pairs to create a Bessel function

Hayashi (2015) presented a visual for how SPAC identifies the direction of the wave propagation (Figure 2.1). The figure shows how these microtremors, in the form of surface waves, can propagate across an array (Figure 2.1a), and how the coherence of these waves can be plotted for each receiver paired with the central receiver (Figure 2.1b). This coherence between receiver pairs takes the form of a cosine function, and the cosine functions can be averaged to create a Bessel function (Aki, 1957) as shown in Figure 2.1c. This Bessel function is used to identify the velocity of the waves through the spatial correlations found between the array's receivers – hence the term spatial autocorrelation.

In mathematical terms, Malagnini et al. (1993) summarize Aki's procedure succinctly. They show that, given a circular array with a receiver in the center, Aki defined the spatial correlation function as;

$$\phi(r,\lambda) = \langle u(x,y,t) \cdot u(x + r\cos\lambda\,, y + r\sin\lambda\,, t)\rangle,$$

where $u(x,y,t)$ is the waveform/velocity observed at point (x,y) at time t, r is the radius of the circular array, λ is the azimuth between each receiver and the central receiver, and $\langle \cdot \rangle$ indicates the azimuthal average. This azimuthal average (the average of the functions between each receiver and the central receiver) can be shown to equal;

$$\phi(r) = \frac{1}{\pi}\int_0^{\pi} \phi(r,\lambda)d\lambda.$$

Aki (1957) showed that this function can be related to the power spectrum, $\Phi(\omega)$, by the zeroth-order Hankel transform to give the equation;

$$\phi(r) = \frac{1}{\pi}\int_0^{\infty} \Phi(\omega) J_o\left(\frac{\omega r}{c(\omega)}\right) d\omega,$$

where ω is the angular frequency, $J_o\left(\frac{\omega r}{c(\omega)}\right)$ is the zeroth-order Bessel function of the first kind, and $c(\omega)$ is the phase velocity of the wave which varies with frequency. Note that dispersive waves are characterized by two types of velocity – phase velocity and group velocity (Rayleigh, 1877). The phase velocity is the speed of a single phase of the waveform, whereas the group velocity is the velocity of a packet of "group" waves. Think of the waves of a ripple in a pond; the waves move up and down at a certain speed but the packet of waves disperse at another speed. Phase velocities are the ones used in SPAC; group velocities are rarely used for near-surface applications (Foti et al., 2014). Aki goes on to show that the phase velocity can be found if the recordings are bandpass filtered over a frequency range centered on the fundamental frequency, ω_o, and the spatial correlation function is normalized to the power spectrum which gives the equation;

$$\phi(r, \omega_o) = J_o\left(\frac{\omega_o r}{c(\omega_o)}\right).$$

Later, Bracewell (1978) used similar logic and formulated equations to find a wave's phase velocity from an array without a central receiver.

2.1.2 Acquisition Equipment

The receivers used in this study were 4.5 Hz geophones which are a form of velocimeter (i.e. electrodynamic velocity transducer). They have the ability to record very small displacements that, in this study, come from vibrations in the soil. They are typically of the moving coil type which houses a coil that sits in a magnetic field. This coil is suspended by a spring(s) which encompasses the magnet that produces the magnetic field. The magnet is permanently fastened to the casing so that, when the

geophone moves, the relative movement of the coil creates a small voltage that is proportional to the relative velocity of the geophone (Foti et al., 2014). A visual of the type of geophone used in this study and a cross-section of a typical moving coil type geophone can be seen in Figure 2.2.

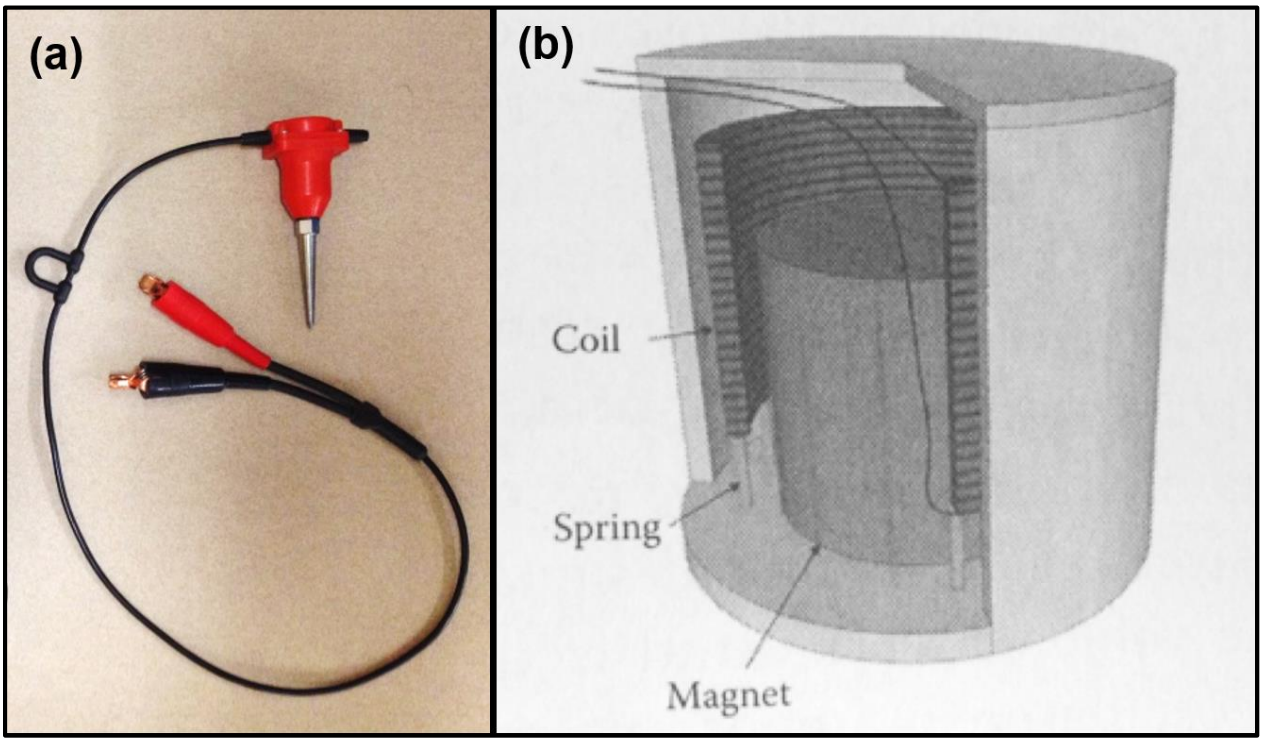


Figure 2.2 Geophone visual; (a) 4.5 Hz geophone used in this study and (b) cross-section of a typical moving coil type geophone (Foti et al., 2014, p. 193)

The geophone sends the analog information to the data acquisition (DAQ) system which digitizes the signal. As well as conditioning the data, the DAQ system also directs the sampling rate which needs to be high enough to avoid aliasing. Aliasing can create distortions and artifacts in the data that can lead to erroneous analysis. However, the Nyquist-Shannon theorem says that no information will be lost by regular sampling if the sampling frequency is two times greater than the highest frequency of the sampled signal (Telford and Geldart, 1990). For example, this study used a sampling frequency of 2

milliseconds which corresponds to 500 Hz, so no information will be lost for wave frequencies that are less than or equal to 250 Hz.

2.1.3 Surface Wave Measurement and Analysis

Rayleigh waves are the type of surface wave analyzed in this study. Other studies have used Love waves, but their horizontal particle motion requires horizontally-oriented receivers and Love waves will not form if a low-velocity layer exists below a higher-velocity layer (Hudson, 1980). In heterogeneous material, both types of waves are geometrically dispersive, meaning that their velocity depends on the wave's frequency (Aki and Richards, 2002). Different frequencies will sample different depths of the geologic material. Higher-frequency waves sample shallower material and lower-frequency waves sample deeper material. This can be more easily explained in terms of wavelength (which is inversely proportional to frequency); waves with larger wavelengths will reach deeper into the surface and will sample more material than waves that have shorter wavelengths. An illustration of this can be seen in Figure 2.3.

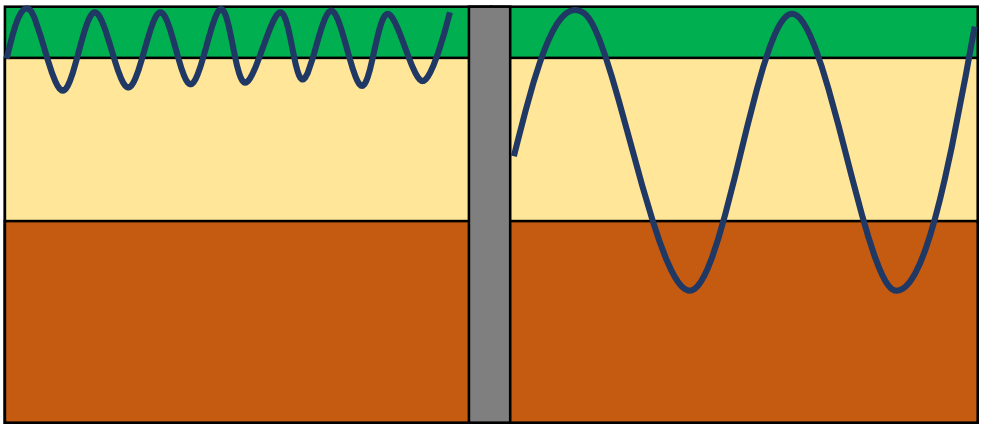


Figure 2.3 Illustration showing how shorter wavelengths (left) sample shallower material and longer wavelengths (right) sample deeper material

Waves with large wavelengths/low frequencies tend to have higher velocity values since deeper material tends to be stiffer than shallower material. For passive waves, the wave source is not controlled, but it is assumed that a broad frequency range will come from the ambient noise. Waves with frequency ranges below 1 Hz are mainly generated from global, geophysical events (ocean waves in particular), and waves with frequency ranges above 1 Hz are mainly generated by human events such as traffic and industrial activities (Foti et al., 2014). As a rule of thumb, the particle motion from a wave is confined to one wavelength from the free surface (Achenbach, 1973). So wavelengths that are smaller than the thickness of the first layer will only sample material from that first layer, but larger wavelengths will sample multiple layers and the velocity of these large-wavelength waves will be governed by some combination of the properties within each layer.

The array diameter is one of the main factors that controls what size waves will be recorded and, therefore, how deep the array will sample. There are different recommendations for how large to make an array if a certain wavelength is desired. Tokimatsu (1995) recommends an array diameter that is one-third of the desired wavelength and studies like Asten and Henstridge (1984) use an array diameter equal to the desired wavelength. But even when a wavelength is recorded, the velocity values are only accurate to a certain depth. Through analysis of variability, Rix and Leipski (1991) found that velocities could be accurately estimated at a depth equal to half of the wavelength. Other studies have recommended different fractions of the wavelength (e.g. Joyner et al., 1981), but Rix and Leipski's factor of 0.5 is the one used in this study.

Once the waves are recorded, analysis of these waves is done in the frequency domain since their velocity is frequency dependent. This analysis is usually carried out by generating a dispersion curve (an example of which can be seen in Figure 2.4).

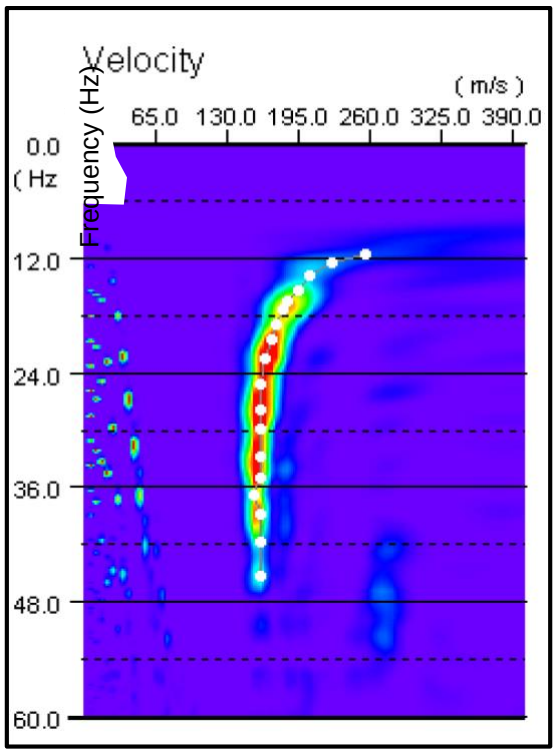


Figure 2.4 Example dispersion curve from the Geogiga Surface Plus (2012) manual

The dispersion curve shows the relationship between the phase velocities and their corresponding frequencies. This curve is generally the goal of surface wave testing because an inversion process can be used to find the shear wave velocity profile, and shear wave velocity is commonly used to predict material properties or behavior (e.g. Andrus and Stokoe, 1996, Tezcan et al., 2006, and Thitimakorn, 2013). The dispersion curve resembles a decreasing exponential curve for sites where the geologic material gets progressively stiffer with depth. This geologic behavior is common, even in homogenous soil, because the weight of the overlying material tends to compress the material below which creates stiffer units and higher velocity values (Santamarina et al., 2001). This

behavior is referred to as being normally dispersive. If the site does not follow normally dispersive behavior, the dispersion curve will depart from the likeness to an exponential curve.

Not only will the dispersion curve change with different material layering, it will also change depending on what mode of propagation is recorded. Figure 2.5 shows an example of how the dispersion curve can change depending on what mode is recorded.

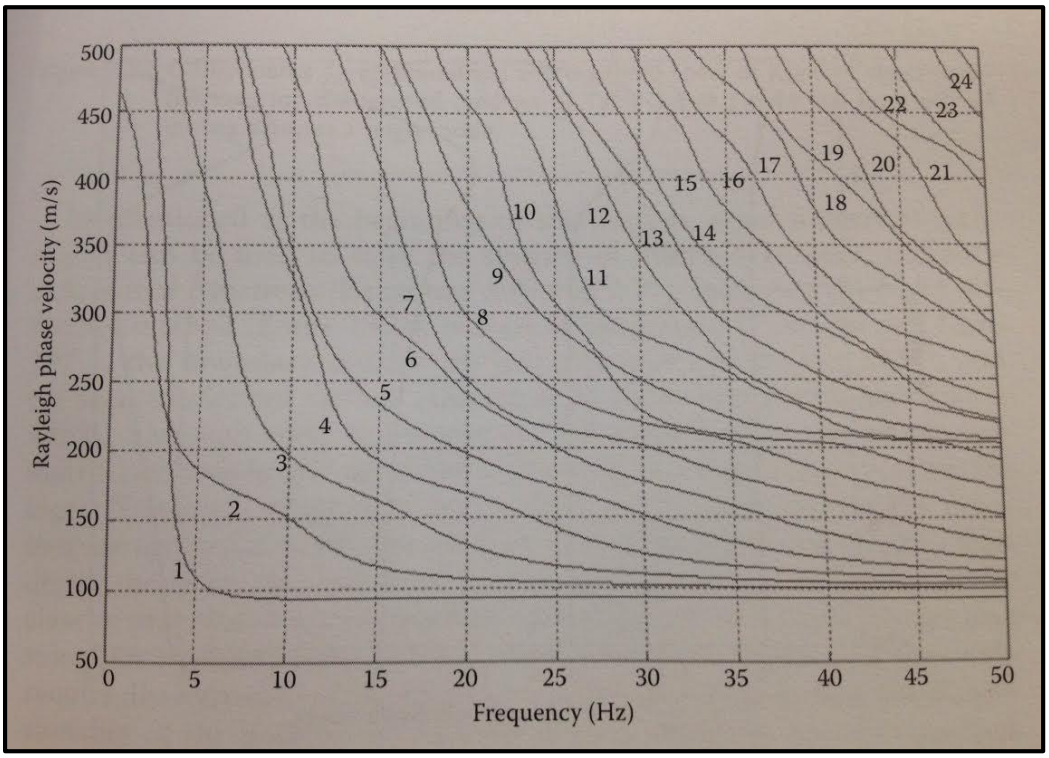


Figure 2.5 An example of how dispersion curves change with differing modes (Foti et al., 2014, p. 75); Each curve represents a single mode of propagation

Multiple modes will sometimes contribute to a dispersion curve, but current methods have difficulty identifying anything but the first mode of propagation (Foti et al., 2014). For normally dispersive sites, the propagation is dominated by the fundamental mode, so the dispersion curve can be assumed to be from the fundamental mode (Lai, 2005). There

can be more contribution from higher modes if a site is inversely dispersive (if high-velocity layers overly low-velocity layers), as seen in Figure 2.6.

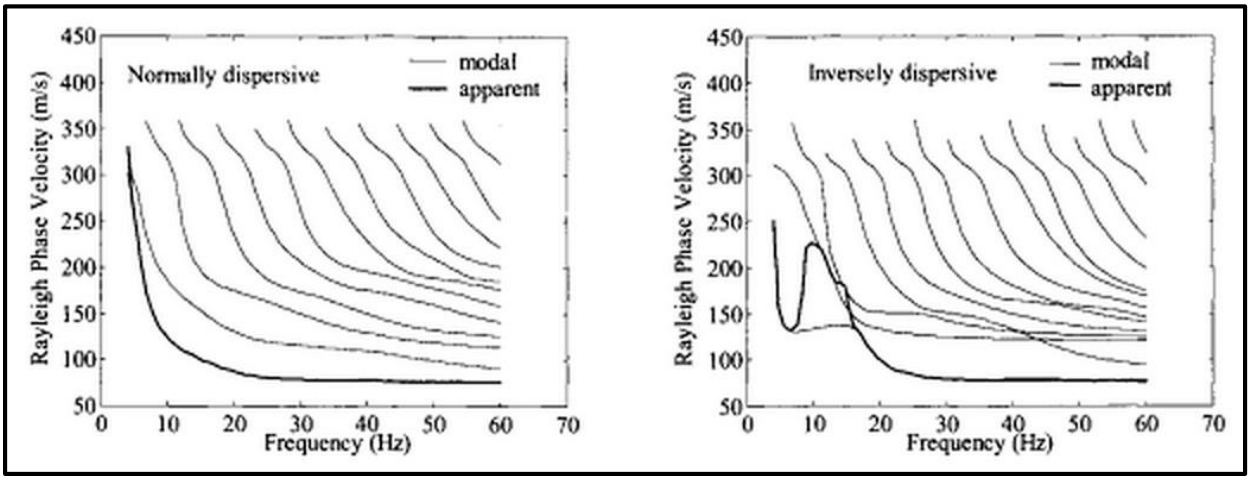


Figure 2.6 Modal and apparent Rayleigh dispersion curves for normally dispersive (left) and inversely dispersive (right) material (Lai, 2005, p. 116)

2.1.4 Inversion

After the dispersion curve has been identified, an inversion process can be used to create a shear wave velocity profile. This inversion process is typically carried out by a computer program (this study used the software Geogiga Surface Plus, 2012). The goal of the process is to create a soil profile that has a theoretical dispersion curve that matches the dispersion curve observed in the field. It does this by assuming a profile and iterating until the theoretical curve aligns with the observed curve. The most common inversion techniques use least squared regression to fit the theoretical dispersion curve to the observed curve (Foti et al., 2014). The residuals from this technique give a misfit error which tells how well the curves match. Through continued iteration, the goal of the process is to minimize the misfit error. The software used to analyze the data in this study (Geogiga Seismic Pro, 2012) uses a genetic algorithm to perform the inversion.

Foti et al. (2014) describe how these inversion processes function. Ordinarily, the inversion of surface wave data is carried out by a local search procedure as opposed to a global search procedure. Most of these local search procedures use calculus-based methods that linearize a non-linear functional (the dispersion curve, in this case) at each iteration until a stationary point is reached. The dispersion curve needs to be sufficiently smooth for this to happen – smooth enough for its Frechet derivatives to exist. The Frechet derivative is outside the scope of this study, but can be seen as a generalization of the gradient to arbitrary vector spaces (Stover and Weisstein). Along with a sufficiently smooth curve, the solution will only converge to one that is similar to the field conditions if the initial estimate is sufficiently close to the solution (Virieux and Operto, 2009). In other words, the velocity model that is initially assumed needs to be similar to the actual model in order to have reliable results. This makes a priori information a valuable asset. Nearby borings, layer thicknesses, and approximate velocities are examples of information that would help constrain the inversion process. Most programs can build an initial model automatically, but this model is built assuming that the profile is normally dispersive (Foti et al., 2014). If the profile is not expected to be normally dispersive, caution should be used when automatically building an initial model as the results will favor a normally dispersive profile.

2.1.5 Limitations

During data acquisition, it is assumed that the waves being recorded are surface waves which are coming from an isotropic wave field and entering the array as a plane. If the waves are recorded too close to the source then body waves could be influencing the data, although this usually poses more problems for active sources than for passive

sources. At 1 to 2 wavelengths from the source, the contribution of body waves becomes negligible and the wave field is dominated by Rayleigh waves (Lamb, 1904). For passive testing with a two-dimensional array, these waves need to be coming from an isotropic wave field (Horike, 1985). This isotropic wave field is required because the SPAC method assumes that random waves are entering the array and are propagating in all directions with equal probability. This ideal situation may be reasonably assumed if there are multiple, spatially distributed sources (Foti et al., 2014). The assumption that the waves are entering the array as a plane is also due to an assumption in the analysis. Rectangular coordinates are used when analyzing surface wave data, but polar coordinates would be needed if the array was too close to the source (Foti et al., 2014). As a rule of thumb, the closest geophone should be 1.5 to 2 wavelengths away from the source in order to avoid this (Foti et al., 2014). Along with these, noise in the recording can also cause large uncertainties when creating the dispersion curve (O'Neill, 2003).

Once the waves are recorded, there are limitations in the inversion processing. The main one being that the inversion process is not unique; a dispersion curve can correspond to many different shear wave velocity profiles (Lai et al., 2005). The only aid to this problem is the use of a priori information. Once the profile is created, the resolution of the model is the other key issue. The more layers a model has, the more uncertainty will be associated with each layer (Parker, 1977). Foti et al. (2014) shows this in Figure 2.7 where the standard deviations for a five-layer model are much smaller than those for a ten-layer model when all other variables are held constant.

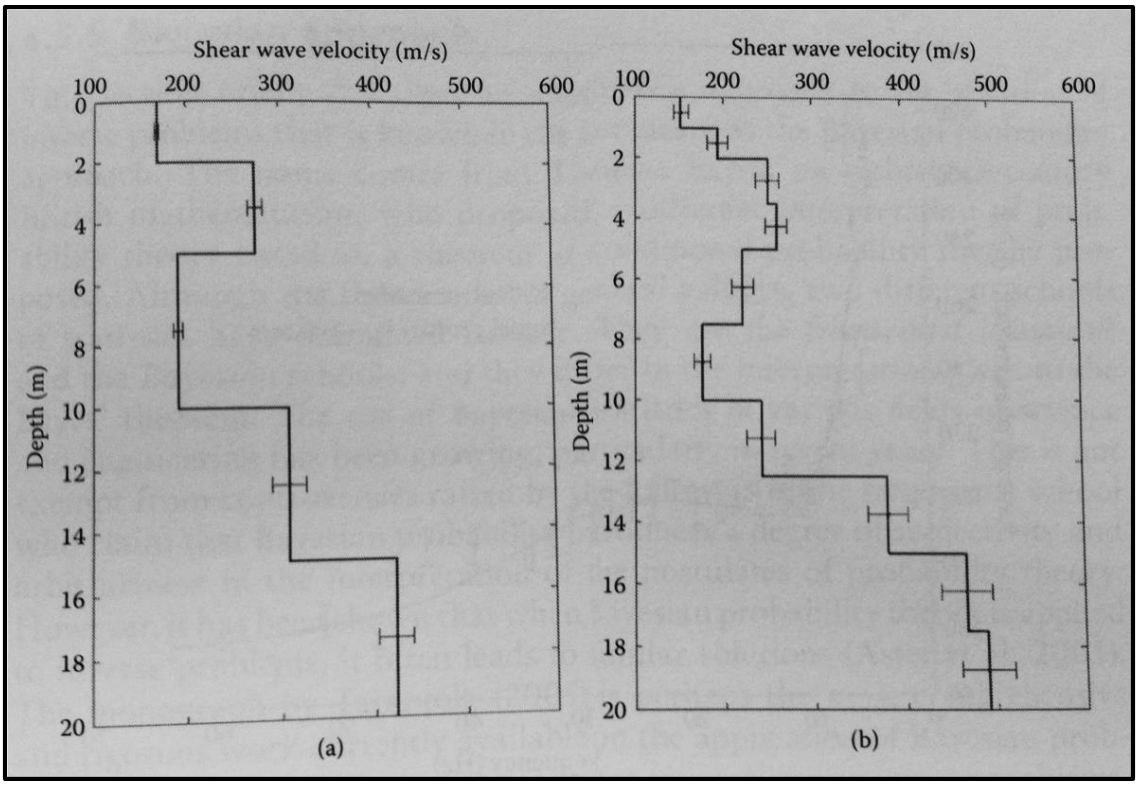


Figure 2.7 Standard deviations for Vs models that only differ by number of layers; (a) five-layer model and (b) ten-layer model (Foti et al., 2014, p. 347)

Notice that the uncertainties at depth, even for the five-layer model, are greater than those at shallower depths. This is common, as shown in research like Tuomi and Hiltenum (1996), Marosi and Hiltunen (2004), and Lai et al (2005). This is because there are, generally, a large amount of rays (wave propagations) that define the shallow layers, but few rays are recorded that define the layers at depth (Foti et al., 2014). An additional factor that causes this loss of resolution is random error which is introduced during the acquisition of the data. Lai et al. (2005) have shown that random errors are present in the dispersion curve values and that these errors cause loss of resolution. They also found that these errors increase with decreasing frequency. All of this makes deep, thin layers very hard to distinguish.

2.2 Review of Geostatistical Methods

Most studies use geostatistical tools to find the variability and the correlation of some material parameter in order to estimate the value of a parameter at unsampled locations. The estimation is most commonly done via kriging. This study uses the tools to find the correlation range – not to estimate unsampled locations. Since the objective of this study is to find a range, estimation is unnecessary (and redundant due to the proximity of the surveys). Because of this, the following review of geostatistics is tailored towards variance rather than estimation.

2.2.1 Introduction to Geostatistics

The goal of geostatistics is to see how spatial data varies at certain distances. This is done by comparing each data point to the others and seeing how similar these data pairs are given their spacing. The comparison can be done graphically; for data separated by a certain lag distance, h, the data pairs can be plotted on a scatterplot. This is called an h-scatterplot, where h is defined as the lag/distance that separates the data pairs. Notice that the values will cluster around a 45^o line if they are similar. On average, most geologic data sets are similar at close lag distances (high values are close to high values and vice versa) and dissimilar at far lag distances (Isaaks and Srivastava, 1989).

Figure 2.8 shows an example of an h-scatterplot for five data points using a lag distance of 5 units where h represents the spacing between the data pairs, $V(x)$ represents the first point of the data pair, and $V(x+h)$ represents the second point of the data pair. Notice how data pairs that have similar values stay close to the diagonal whereas dissimilar pairs plot further from this line. H-scatterplots can be created for additional lag distances to see the similarities or dissimilarities at other separations, and different ways

of quantifying the clustering around the 45^o line will tell how spatially continuous the data is. One way of achieving this is to find the correlation coefficient for each h-scatterplot and to plot this correlation value for each lag distance. An example of the correlation coefficient with lag distance can be seen in Figure 2.9. The plot ranges from 1.0 being perfectly correlated to 0 being uncorrelated, and it shows that data pairs at close separation are similar but those at longer separations are minimally correlated.

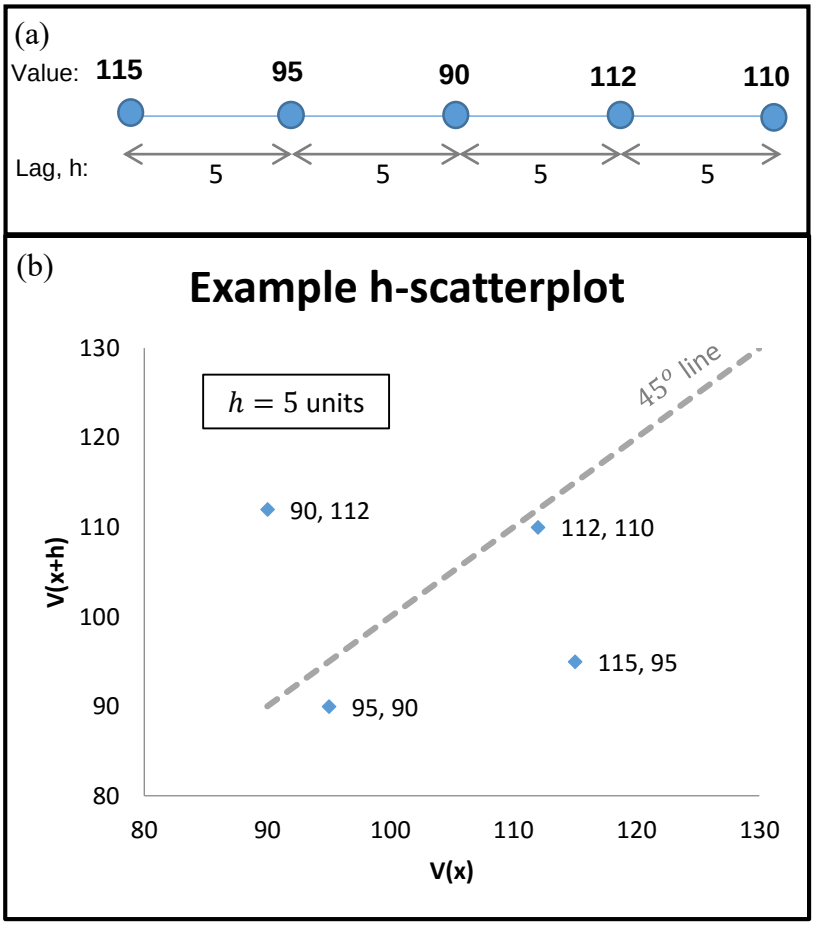


Figure 2.8 Example h-scatterplot; (a) values along a line with 5 unit spacing, and (b) the corresponding h-scatterplot

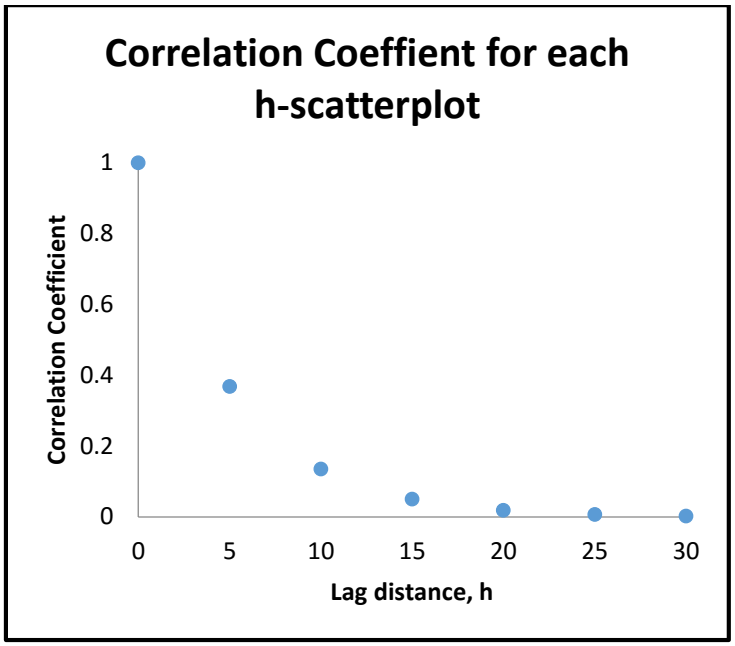


Figure 2.9 Example of how the correlation coefficient varies with lag distance

Other ways of quantifying the h-scatterplot is to find the covariance of the plot. The definition of covariance is similar to that of correlation and, because of this, these plots share the same shape. Another way to model the h-scatterplot data is to find the moment of inertia around the 45^o line as done by Matheron (1965). Since the moment of inertia measures the "fatness" of the cloud of data, this value should increase with increasing lag. This is because nearby data (data with small lag distances) are expected to be similar and as the lag distance gets larger the data pairs are expected to become more and more dissimilar. On the h-scatterplots, this will show points close to the 45^o line at small lag distance and points further from the 45^o line (increasing "fatness") at larger lag distances. This expansion of points will be reflected by an increase in the moment of inertia value. An example of this can be seen in Figure 2.10.

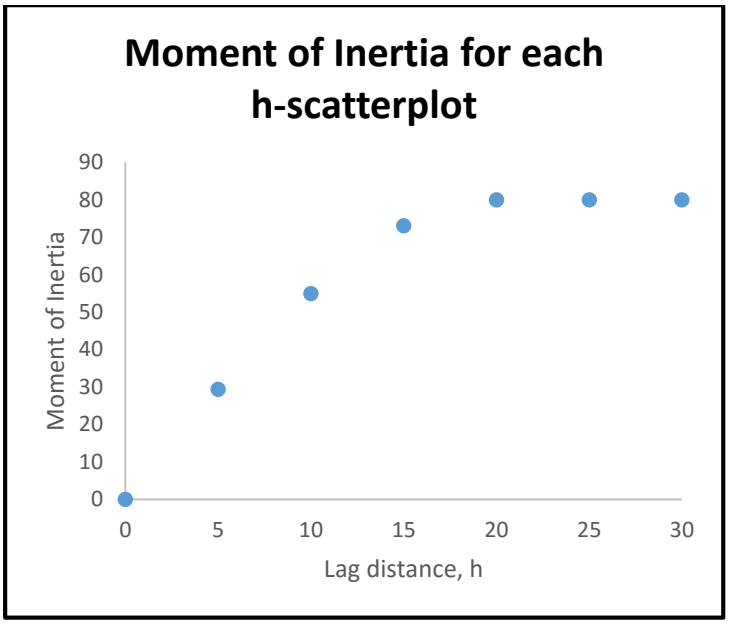


Figure 2.10 Example of how the moment of inertia varies with lag distance

The moment of inertia about the diagonal of the h-scatterplots is defined as $\frac{1}{2n}\sum(x_i - y_i)^2$, where n is the number of data pairs, x_i is the value $V(x)$, and y_i is the value at a certain lag distance $V(x+h)$. The factor of ½ is present because we are interested in the perpendicular distance away from the diagonal line.

The relationship between either the moment of inertia, the covariance, or the correlation coefficient with the lag distance are three common descriptors in geostatistics. The relationships are referred to as the semi-variogram, the covariance function, and the correlogram, respectively.

2.2.2 Spatial Continuity Functions

To save time, it is common to skip plotting each h-scatterplot and instead to graph the semi-variogram, the covariance function, or the correlogram directly. Equations have been formulated to do just this. The semi-variogram can be plotted by using the equation;

$$\gamma(\boldsymbol{h}) = \frac{1}{2N(\boldsymbol{h})} \sum_{(i,j)|h_{ij}=h} \left(v_i - v_j\right)^2$$

where the data values are $v_1, \ldots, v_n$ and the summation is over the $N(\boldsymbol{h})$ data pairs whose locations are separated by h. Notice that h is bolded in the equation because it represents multiple values instead of being a constant. The semi-variogram represents half of the average squared difference between data pairs, and can be shown to equal the variance of the values being estimated (Isaaks and Srivastava, 1989, p. 221-222).

In similar notation, the covariance function can be calculated using;

$$C(\boldsymbol{h}) = \frac{1}{N(\boldsymbol{h})} \sum_{(i,j)|h_{ij}=h} v_i \cdot v_j - m_{-h} \cdot m_{+h}$$

where m_{-h} and m_{+h} are the mean values of all the data points whose locations are $-h$ and $+h$ away from some other data location, respectively. In other words, for each data pair, $-h$ denotes the first point and the $+h$ denotes the second point. These mean values are defined as;

$$m_{-h} = \frac{1}{N(\boldsymbol{h})} \sum_{i|h_{ij}=h} v_i$$

$$m_{+h} = \frac{1}{N(\boldsymbol{h})} \sum_{j|h_{ij}=h} v_j$$

and they typically do not equal each other in practice (Isaaks and Srivastava, 1989).

The correlogram is the same as the covariance function, but is scaled by the standard deviations of the data pairs. It is defined as;

$$\rho(\boldsymbol{h}) = \frac{C(\boldsymbol{h})}{\sigma_{-h} \cdot \sigma_{+h}}$$

where σ_{-h} and σ_{+h} are the standard deviations of all the data values whose locations are $-h$ and $+h$ away from some other data location, respectively. Their equations are;

$$\sigma_{-h}^2 = \frac{1}{N(\boldsymbol{h})} \sum_{i|h_{ij}=h} v_i^2 - m_{-h}^2$$

$$\sigma_{+h}^2 = \frac{1}{N(\boldsymbol{h})} \sum_{j|h_{ij}=h} v_j^2 - m_{+h}^2.$$

The shape of a typical semi-variogram and covariance function is shown in Figure 2.11. Notice that they are the same except for being flipped horizontally (this happens if the data has a constant mean value). The semi-variogram ranges from 0 to its plateau whereas the covariance function starts at the semi-variogram's plateau and decreases to 0. The correlogram matches the shape of the covariance function, but is scaled so that the maximum value is 1.

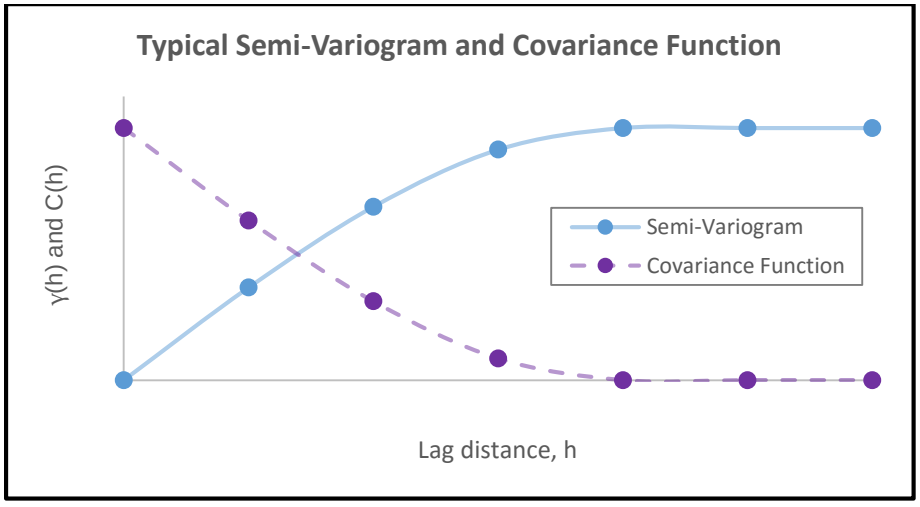


Figure 2.11 Typical shape of the semi-variogram and covariance function

These equations help to quantify how the data changes with distance, but they only allow a description for a discrete amount of lag distances. For example, if there are no data pairs that are close to being separated by 10 units, then no value can be calculated

for this lag. Because of this, the natural progression is to create a model that fits the spatial continuity functions so that all lag values can be analyzed.

2.2.3 Spatial Continuity Models

There are not many earth science applications that are understood well enough to create a deterministic model (Stein and Stein, 2014). Because of this, it is assumed that the process(es) are too complicated to accurately model and they are instead assumed to be random processes. Although this may not be the case (usually it is multiple, complex processes that produce the values), the random assumption tends to give accurate results (Webster, 2000).

For estimation purposes, the models used to describe the spatial continuity need to be positive definite (the eigenvalues need to be positive). The derivation of this condition is outside the scope of this study, but can be referenced in Strang (1980). Among other things, positive definiteness is required because it guarantees that the estimations exist, that they are unique, and that they are stable (Isaaks and Srivastava, 1989).

Bohling (2005) describes how geostaticians use the semi-variogram more than the other spatial continuity functions primarily because it tends to filter out the influence of a spatially varying mean value. The covariance function requires second-order stationarity of the variable (Oliver and Webster, 2015) which means that the random process is assumed to have distributed the values with a constant mean (i.e. the average of the values is the same at every sub-location). This can be seen in the definition of the covariance function – it includes terms for the mean values and these mean values are not a function of x (they are assumed to be constant). This allows the covariance function to rely solely on the lag distance and not on location. The semi-variogram only requires that

the first differences, $V(x) - V(x + \boldsymbol{h})$, are second-order stationary (Bohling, 2005) which is what Matheron (1963) referred to as intrinsic stationarity. This can also be seen in the definition of the semi-variogram; the only variables involved in the summation are the data pairs. Oliver and Webster (2015) describe how intrinsic stationarity assumes the expected differences of the values to equal zero;

$$E[V(x) - V(x + \boldsymbol{h})] = 0$$

which replaces the covariance of the residuals with the variance of the differences;

$$\text{var}[V(x) - V(x + \boldsymbol{h})] = E[\{V(x) - V(x + \boldsymbol{h})\}^2] = 2\gamma(\boldsymbol{h}).$$

And this spatial relationship defines the variogram (half of this value defines the semi-variogram). This is a useful replacement because it relieves the requirement for a constant mean value. Since the semi-variogram does not rely on a constant mean value, it can be defined in some cases where the covariance function cannot be defined, particularly if the semi-variogram increases without bound (Bohling, 2005).

There are many models that could be used to define the spatial continuity, but some "basic models" have been created that are positive definite and vary enough to satisfactorily fit most semi-variograms that are likely to be encountered (Isaaks and Srivastava, 1989). There are two, general types of models: ones that reach a plateau and ones that do not. This plateau is called the sill (denoted as C) and the lag distance where the model reaches the sill is called the range (denoted as a). The models that reach a sill are called transition models, and the basic transition models are the spherical model, the exponential model, and the Gaussian model (Isaaks and Srivastava, 1989). These models can be seen in Figure 2.12.

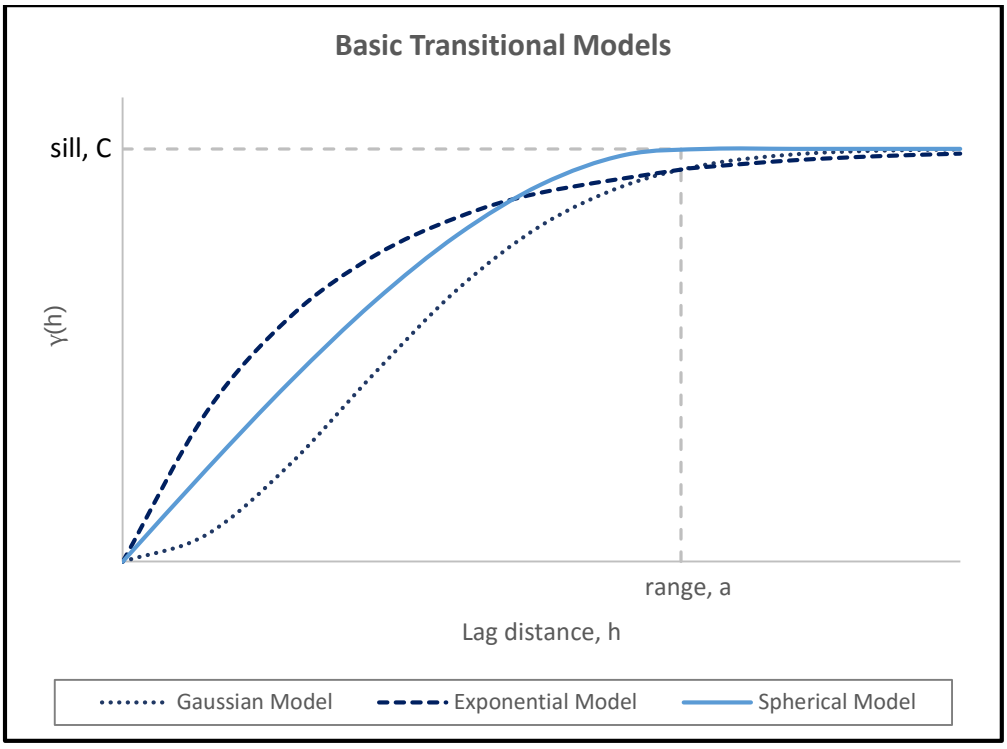


Figure 2.12 Basic transition models

The equation for the spherical model is;

$$\gamma(\boldsymbol{h}) = \begin{cases} C\left(\dfrac{3\boldsymbol{h}}{2a} - \dfrac{\boldsymbol{h}^3}{2a^3}\right) & \text{when } \boldsymbol{h} \leq a \\ C & \text{when } \boldsymbol{h} > a \end{cases}$$

the one for the exponential model is;

$$\gamma(\boldsymbol{h}) = C\left[1 - \exp\left(-\frac{3\boldsymbol{h}}{a}\right)\right]$$

and the one for the Gaussian model is;

$$\gamma(\boldsymbol{h}) = 1 - \exp\left(-\frac{3\boldsymbol{h}^2}{a^2}\right).$$

A function that is second-order stationary will reach a sill and this sill defines the a priori variance, σ^2, of the random function that is assumed to have created the values

(Oliver and Webster, 2015). In other words, it represents the variance for the distribution of values at the site. And the lag distance that corresponds to the sill (the range) defines the limit of spatial correlation. Values with lag distances greater than this are spatially uncorrelated or independent (Oliver and Webster, 2015). Notice that the exponential model and Gaussian model only reach the sill asymptotically. These models do not technically have a range since the model never reaches the sill, but a practical/effective range is used in practice that is defined as the lag distance at 95% of the sill value (Isaaks and Srivastava, 1989).

Unbounded models are those that do not reach a sill, and basic unbounded models are the power model and the linear model (seen in Figure 2.13). The equation for the power model is defined as;

$$\gamma(\boldsymbol{h}) = g\boldsymbol{h}^{\beta} \quad \text{for } 0 < \beta < 2$$

where g describes the intensity of the variation and β describes the curvature. The limits 0 and 2 are excluded because $\beta = 0$ creates constant variance for all lag distances and $\beta = 2$ creates a parabolic function which means that the process is not random (Oliver and Webster, 2015). If $\beta = 1$, the model becomes a linear model with g representing the slope of the line. Because the models are unbounded, the covariance function and correlogram do not exist, but the semi-variogram does exist and fulfills Matheron's (1965) intrinsic hypothesis. This is the intrinsic stationarity discussed earlier in the chapter where the semi-variogram is not restricted by a constant mean value.

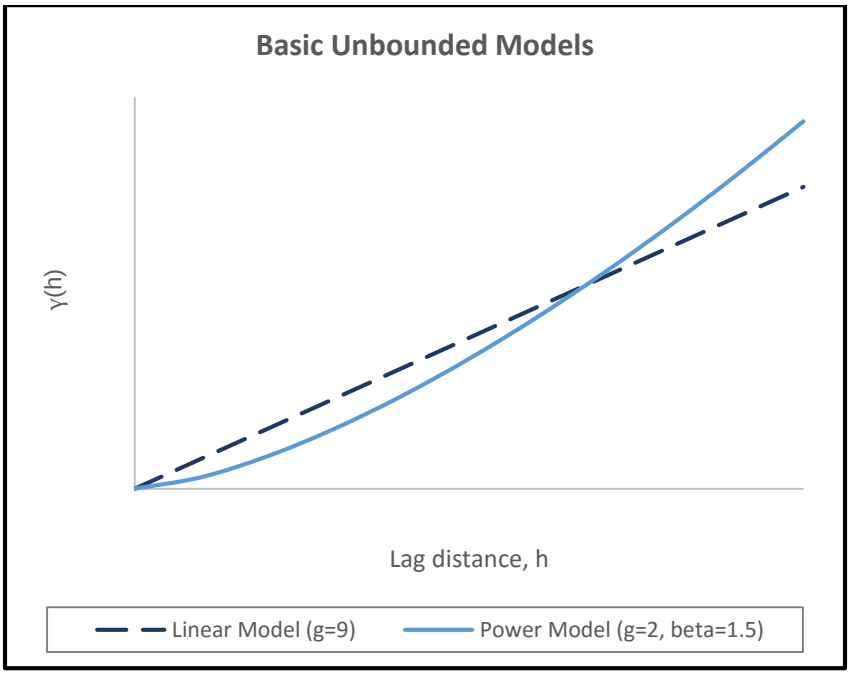


Figure 2.13 Example of basic unbounded models

All of the models shown thus far have intersected at the origin. Although the semi-variogram value for $h = 0$ is strictly 0, values at very small lag distances may be larger than 0 which creates a discontinuity. This implies that neighboring data have an average variance that is significantly larger than 0. This can be from several factors; some examples being random error, variation over distances less than the sampling interval, and measurement error (Oliver and Webster, 2015) which all show the unpredictable component of the values. This jump is called the nugget effect. Oliver and Webster (2015) describe that the term "nugget" came from gold mining because gold nuggets appeared to occur at random and independently of one another. They created an uncorrelated component because the gold content did not relate to the neighboring sites. This effect can be modeled by adding $\omega_o \gamma_o(\boldsymbol{h})$ to the model where ω_o is the height of the discontinuity and $\gamma_o(\boldsymbol{h}) = \begin{cases} 0 & \text{if } h = 0 \\ 1 & \text{otherwise} \end{cases}$. This translates to being 0 at the origin and a constant value otherwise.

When deciding which model to fit, it is best to use the one that gives the smallest sum of squared residuals (Oliver and Webster, 2015). If the sample semi-variogram appears to have a sill, then the behavior of the values near the origin can help select a model as well. If the points near the origin appear parabolic, this resembles the Gaussian model. If a line drawn from the first few points intersects the sill value at two-thirds of the range, this resembles a spherical model. And if a line drawn from the first few points intersects the sill value at one-fifth of the range, this resembles an exponential model (Isaaks and Srivastava, 1989). An example of this can be seen in Figure 2.14.

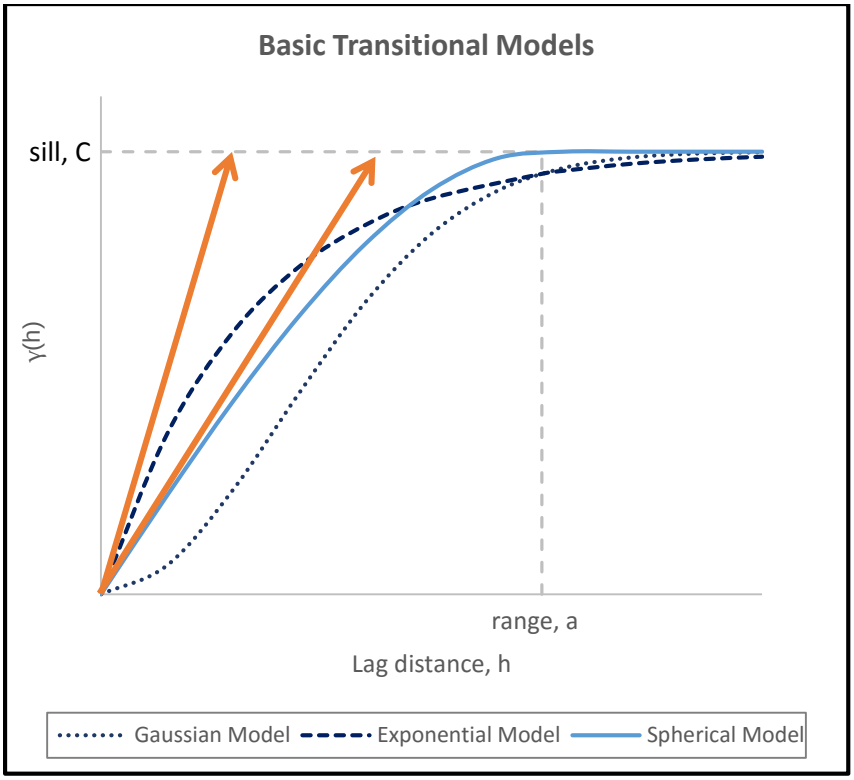


Figure 2.14 Deciding on a model from the first few points in the semi-variogram

2.2.4 Limitations

A major limitation is the presence of a trend in the data. The process that created the values is assumed to be random, but a trend indicates a non-randomness that will give inaccurate values if not accounted for (Oliver and Webster, 2015). This violates the

stationarity assumption because the mean value varies as a function of distance. Instead of the variance being estimated from $V(x) = \mu + \varepsilon(x)$, a trend implies that $V(x) = \mu(x) + \varepsilon(x)$ where $V(x)$ are the values of interest, μ is the mean of the process, and $\varepsilon(x)$ is the spatially-correlated, random residual (Oliver and Webster, 2015). Trends may be fairly obvious if the values are plotted, but a hint that one may be present is if the semi-variogram steadily increases without bound which creates negative correlation between variables separated by large lags (Bohling, 2005). If the trend is defined and a model is fit to the data, the trend can be subtracted from the data and spatial continuity analysis can be completed on the residuals (Bohling, 2005). Data sampled in small ranges (around 20 meters) is usually safe to assume that no prevailing trend will be present (Clark, 1979).

Additional limitations that could affect accuracy is the sample size and the sample interval. If too few data is sampled, the average values of the experimental semi-variogram are less likely to align with the true semi-variogram and there will be more variation in these semi-variogram values. This was studied by Webster and Oliver (1992) and they concluded that a survey should aim for 150 data points with 100 data being a minimum. The sampling interval is important because, if too large, it may not capture enough points inside the range of the model. If the data plateaus, this would lead to a model appearing to be a pure nugget effect (a flat line because there are no data pairs to describe the model at smaller lag distances). Oliver and Webster (2015) recommend a sample interval that allows at least five estimates of $\gamma(\boldsymbol{h})$ before the range. Viewing semi-variogram models of similar sites can help to estimate an approximate range so that the sample interval can be estimated.

Chapter 3: Testing Methods

The particle motion from propagating seismic waves is the information we want to capture. This data is collected using a set of receivers and an acquisition system, and multiple receivers allow the wave propagation to be recorded in both time and space. These recordings can then be used to estimate the properties of the underlying geologic material. The following sections outline what equipment is needed, how the equipment is setup, how to acquire the data, how to process the data, and then how to model the data.

3.1 Equipment

The following list details the equipment needed for passive surface wave testing, and Figure 3.1 illustrates the equipment that this study used while testing (with numbers corresponding to the list below).

1. Laptop with Vibrascope and Geogiga Seismic Pro installed
2. Seismic recording system (DAQ Link III Seismograph)
3. 12V battery for data acquisition (DAQ) system
4. Ethernet cable
5. DAQ power cable
6. 12-channel geophone cable
7. 4.5 Hz, single-channel, vertical-motion geophones (× 12)
8. Measuring tape (minimum of 10 meters)
9. Metal stake

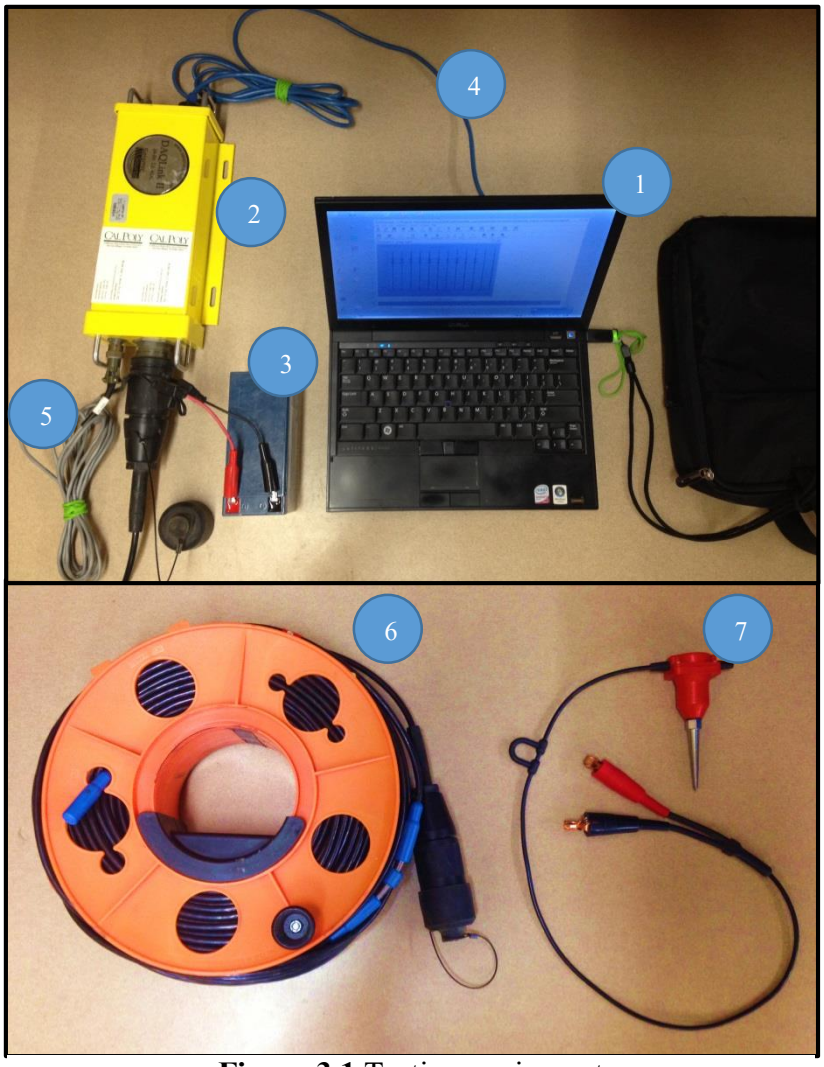


Figure 3.1 Testing equipment

It should be noted that the receivers have inherent limitations. The receivers in the array can only accurately record waves that are above their natural frequency (Lai, 2005) which is the resonant frequency of the oscillator. The amplitude of the waves will be greatly attenuated in the recordings for frequencies below this natural frequency. Since this study uses 4.5-Hz geophones, any wave with a frequency lower than 4.5 Hz will not be accurately measured.

3.2 Field Setup

Although there are many different array types, this study uses a 10-meter-diameter, circular array with 12 geophones. The reason for using this type of array is empirical; past experience and trial-and-error have shown that, compared to other array types, this geometry gives adequate resolution and depth. The performance of an array depends on the array geometry and the properties of the wave field, but there is no universal agreement on the array design (Foti et al., 2014).

This study set all the 10-meter, circular arrays along a straight line with each consecutive array overlapped by 5 meters. The approximate locations of each of these arrays can be seen in Figure 3.2. Every fifth array in the figure is highlighted to reduce clutter and there is a gap of three arrays due to a hill on the athletic field. Each array was labeled based on its distance from a 4-foot, metal stake on the site. This stake is at the center of the fifth array and is labeled in the figure.

Setting up in the field begins by determining the line (array azimuth) and a starting location. The array azimuth in Figure 3.2 was chosen based on the geology, topography, and constraints of the site. In choosing the azimuth, the objective was to create a long array on flat ground that sampled similar material. The starting/reference location was chosen based on logistics and ease of access at the site.

Figure 3.2 This study's array line at Cuesta athletic fields

Since the array is circular, it was easiest to anchor the measuring tape to the center of the array with a stake and rotate radially around the stake to create a perfect circle. The approach we used to place the geophones was to start with the cardinal points of the circle. Using the measuring tape, 4 geophones were placed at each cardinal point – making sure that two of the cardinal points were in line with the array azimuth. Since there are 12 total geophones, the four quadrants that were just created needed to be divided into thirds in order to accommodate the remaining 8 geophones. Estimating these locations visually or measuring the distance by pacing around the circle proved to be sufficient for this study. Each geophone was pushed into the ground so that it had intimate contact with the soil. If the top soil was too stiff to place the geophones by hand, we would hammer a stake partway into the soil to create a void for the geophone spike. This was sufficient because ground coupling and the tilt of the geophones only minimally influence the data (O'Neill, 2003). Once all the geophones were set, the geophone cable was laid around the circle and the geophones were attached to their respective leads. Since we had multiple arrays, we made sure to keep the geophone numbering in the same orientation (e.g. the first and seventh geophones were always along the array azimuth). The cable was plugged into the seismic recording system (DAQ Link III seismograph), and this system was attached to both the 12V battery and the laptop via the power cable and the Ethernet cable, respectively. After these steps, the array is ready for data collection in Vibrascope. An example of what the finished array should resemble can be seen in Figure 3.3.

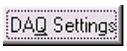
Figure 3.3 A visual of this study's array setup

3.3 Acquisition of Experimental Data

In Vibrascope, it is best to create a new project for each array location so that organization and accessibility of the files remains simple. This can be done by selecting “New Project” from the “File” drop-down menu. In this “New Project” window, as seen in Figure 3.4, input the project name and the project location. Vibrascope will create a new folder with the specified project name and the folder will be placed in the specified project location.

Figure 3.4 New Project window in Vibrascope

Once the new project has been created, the recording parameters can be verified and adjusted in DAQ Setting. This can be found by selecting the "Options" drop-down menu and clicking "Device", or by pressing the "Device" icon 

in the ribbon. In this DAQ Setup window, press 

or press "Q" on the keyboard. This will bring up the "Configuration" window as seen in Figure 3.5. In the acquisition tab of this window, be sure that all 12 channels are selected in the "On" column, that each channel has the "Type" as generic, the "Gain" as 1, the "DC Removal" as Auto at 0.0, the "Units" as V, and the "Scale" as 1.

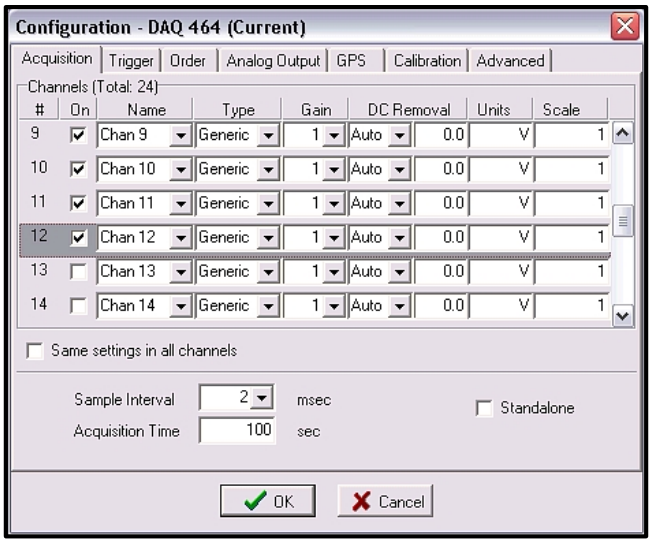


Figure 3.5 Configuration window on the Acquisition tab

For surface wave recordings intended for SPAC or f-k analysis, the "Sample Interval" should be set to 2 milliseconds and the "Acquisition Time" should be set to 100 seconds. The 2 millisecond acquisition time is because analysis is done in the frequency domain and the frequencies of interest are usually below 100 Hz (Foti et al., 2014). The 100-second cap is because, at the 2 millisecond sampling rate, this is the largest file that most software can handle. In the trigger tab on the Configuration window, be sure that "Auto Trigger" is selected. Once this is complete, press OK on the Configuration window and then on the DAQ Setup window to accept the settings and return to the main screen. Note that if the Configuration window did not resemble Figure 3.5, but rather resembles Figure 3.6, then "Advanced Mode" has been turned off.

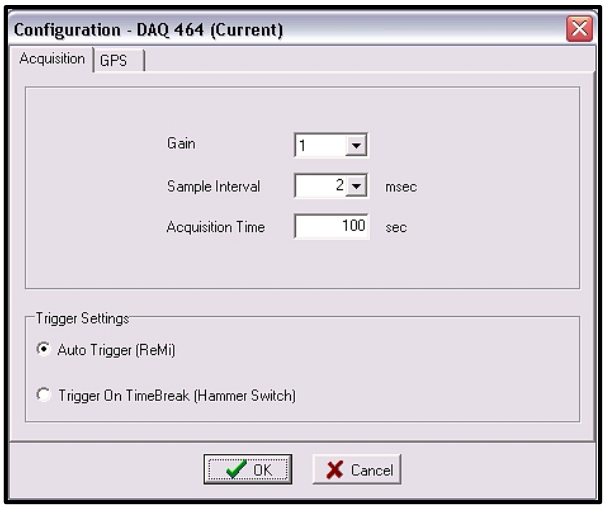


Figure 3.6 The Configuration window if "Advanced Mode" is not selected

To select “Advanced Mode”, drop down the “Options” menu on the main screen and choose “Preferences.” In the Preferences window, make sure that “Advanced Mode” is checked.

When the DAQ settings are correct, the data can be collected by pressing 

in the ribbon or by pressing “A” on the keyboard. It may prove easier to run a 10-second recording before running the first real acquisition. This can be done by changing the acquisition time to 10 in the DAQ Settings window. This 10-second recording should not be saved, but it allows the user to see if all the geophones are receiving data and if the seismic plot looks appropriate without having to wait the full 100 seconds. An example of an appropriate 10-second recording can be seen in Figure 3.7.

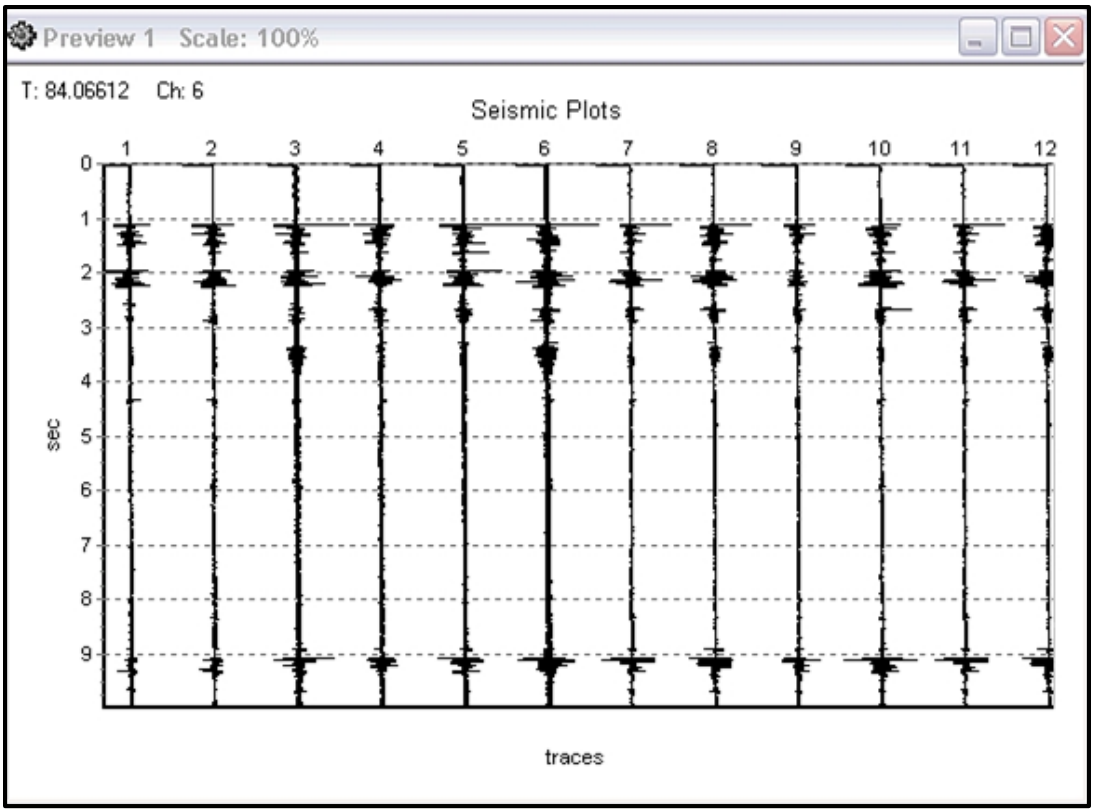


Figure 3.7 An example of an appropriate 10-second recording

If the 10-second recording seems reasonable, then the acquisition time can be changed back to 100 seconds and the first real recording can be taken. Notice that, along the bottom of the main screen, Vibrascope shows how many seconds that the recording has been receiving data after pressing start (circled in Figure 3.8).

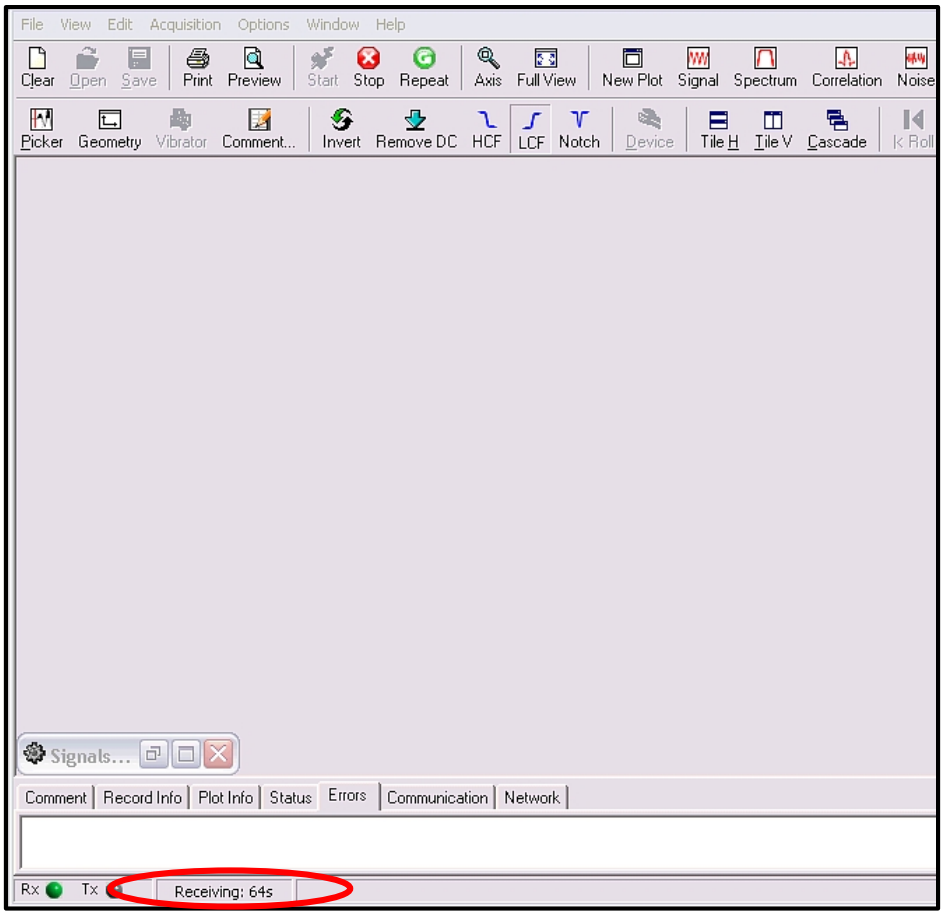


Figure 3.8 Vibrascope's indication of how long the recording has been receiving data

This can be helpful in keeping track of the time and in noting how certain seismic sources affected the data (e.g. if a concrete truck drives by the array at 40 seconds of recording, the seismic plot should reflect this). Once the 100 seconds of recording are complete, a window displaying the signal traces for each geophone should appear on the screen. It is good to check this window for anomalies, but be advised that the traces in the seismic plot are scaled to the largest amplitude that was recorded. This means that if the array recorded a wave with significantly higher amplitude (maybe from a heavy object dropping nearby), then the scale would be adjusted to this amplitude in the seismic plot and the motion of the lower amplitude waves may appear to approach zero. If there was adequate noise during the recording though,

this scaling would just be hiding the lower amplitude waves and the use of the recording would still be appropriate.

If the recording seems suitable, then the next step is to save the recording by selecting "Save" in the "File" drop-down menu or by pressing Ctrl+S. Notice that the name on the banner of the seismic-plot window changes to reflect how many recordings have been saved in the project folder. After saving, the file can be exported into the project folder by selecting "Export" in the "File" drop-down menu. In order for the signal processing software (Geogiga Seismic Pro, 2012) to be able to read these files, they must be saved as SEG-Y Tape Format (*.sgy). Naming the file with some reference to the location makes it easier to organize and identify the files during processing. The exporting window described above can be seen in Figure 3.9.

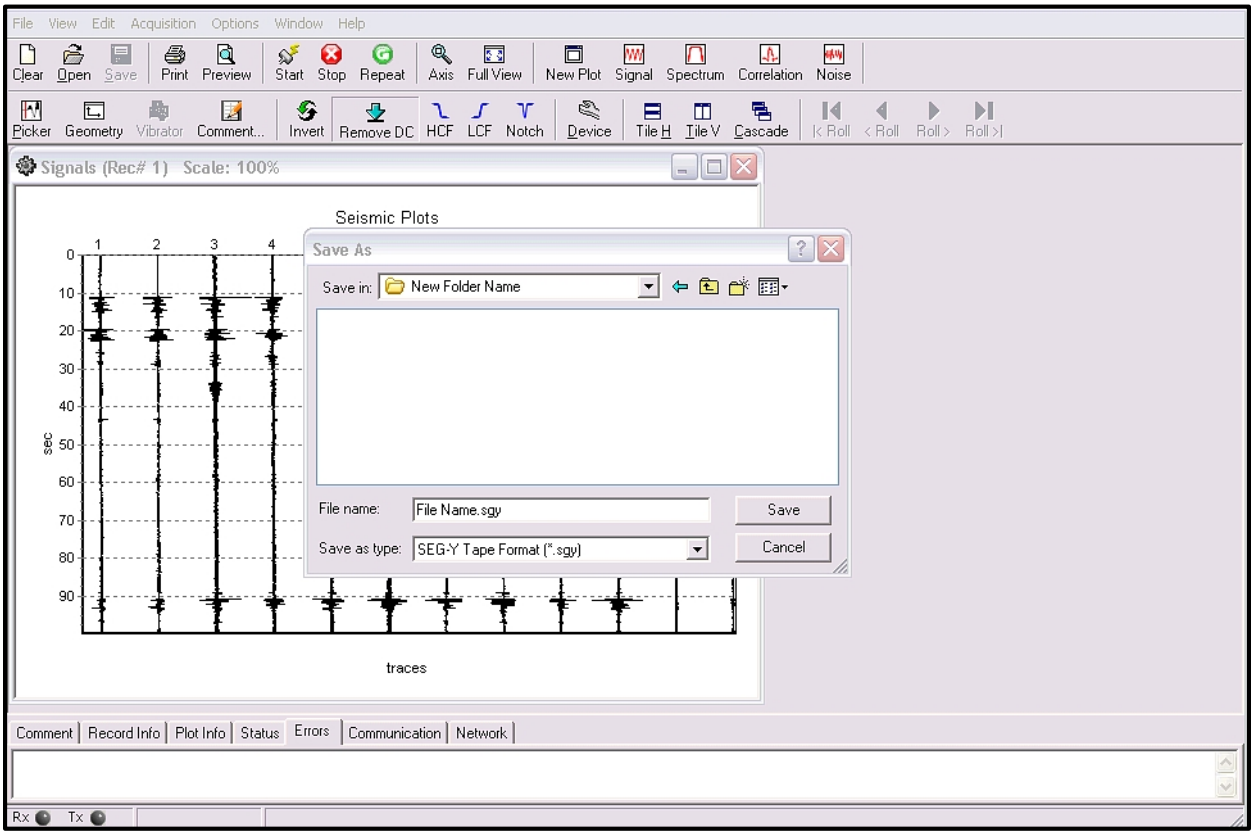


Figure 3.9 Vibrascope's export window

Once exported, complete two more 100-second recordings by following the steps outlined above and export these into the same project folder. 100-second recordings are the largest file size that most software can handle, but more recordings are necessary to fully sample the passive/ambient noise. Three, 100-second recordings will commonly produce sufficiently accurate results. After three recordings have been completed, the array can be moved to a new location and a new project folder can be started.

3.4 Signal Processing

The three recordings at each location can be processed using the Surface Plus program in Geogiga Seismic Pro. Import the files by selecting "Import Seismic" from the "File" drop-down menu. Make sure that "Passive (2D Array)" is selected at the top of this window as seen in Figure 3.10, and then add the files by selecting "Add…" near the bottom of the window (as opposed to the "Add" near the Geometry Definition Group section). The file will be in the project folder that was created in Vibrascope. Highlight all three recordings and press "Open." Notice that the last digit in the "Shot No." column corresponds to which file was recorded first, second, and third.

Once the three files are loaded, the geometry for *each* file needs to be added. This can be done by selecting the first file and then selecting 

which will bring up a "Geometry" window (seen in Figure 3.11). If there are not three columns (X, Y, and Z) in the "Receiver Locations" section then it is likely that "Passive (2D Array)" was not selected in the importing window.

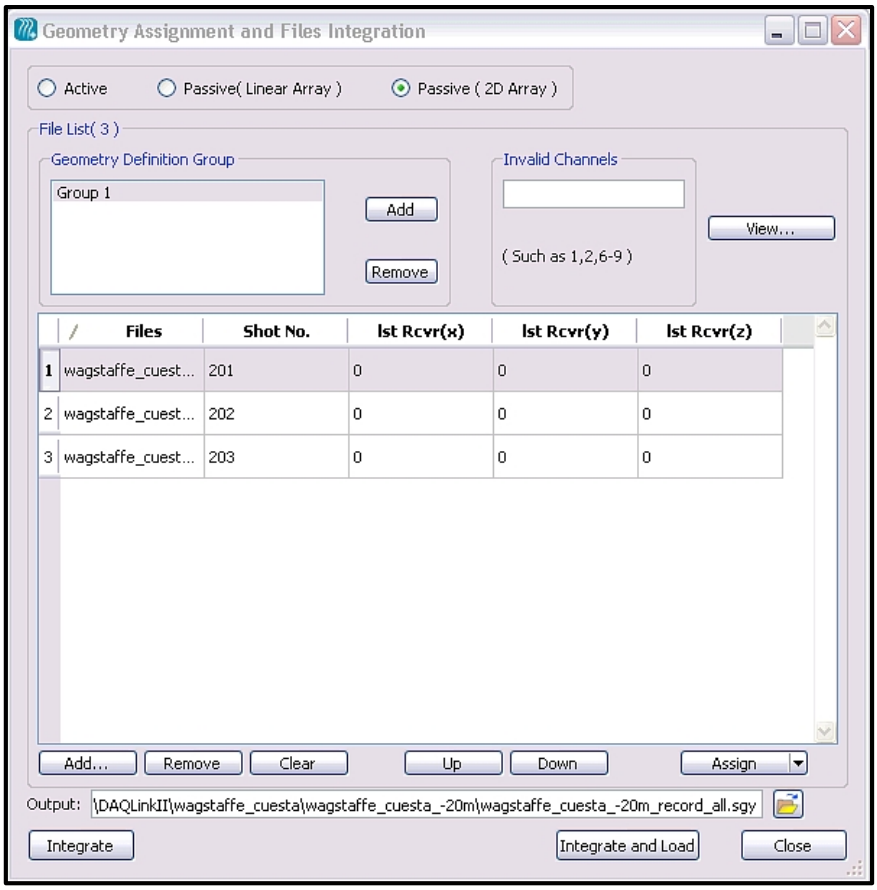


Figure 3.10 Importing seismic data in Geogiga Seismic Pro

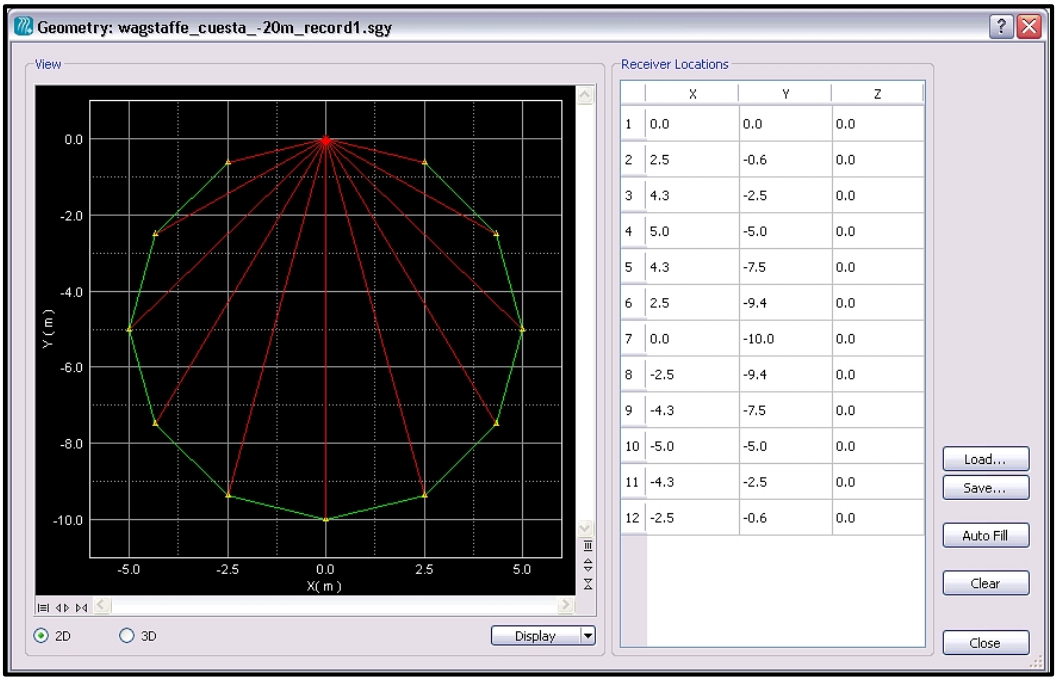


Figure 3.11 Setting geometry in Geogiga Seismic Pro

To load the geometry, press Load... and then double-click the text file (.txt) that lists the receiver locations. It is best to save a copy of this file to the folder where the new Vibrascope projects are being saved. This will make locating the file much easier. The geometry text file for a 10-meter-diameter, circular array with 12 geophones can be seen in Figure 3.12.

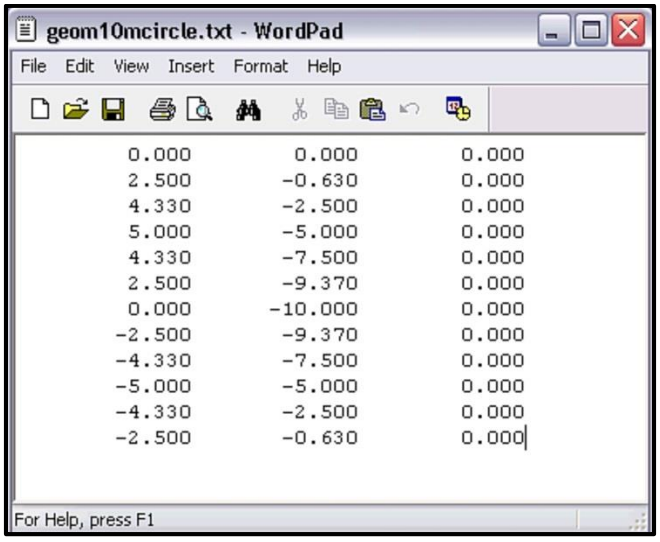


Figure 3.12 Text file with receiver locations

Double-clicking the text file will input the values into the receiver locations and the lines that connect all geophones to the first geophone should turn red. After this, press Close to exit the window. Now click on the next file and repeat the steps above to assign it a geometry. Note that simply using the arrow keys to highlight the next file will not apply the geometry to the correct recording. Also be aware that the "View…" window for the second and third recordings will show a diagram of the circular array with light blue lines, but the geometry will not actually be loaded until the text file is selected. To ensure that the geometries have been applied, make sure that the receiver locations have changed to the correct values and that the lines connecting the geophones to the first geophone have changed from light blue to red. Once all of the files have the correct geometry applied, click Integrate and Load (an error message will appear if all the

geometries have not been applied). Geogiga will link to the “Browse Traces” option and the traces from each recording should be visible on the main screen.

The dispersion settings now need to be changed, but this cannot be done while browsing the traces; one of the traces needs to be selected so that Geogiga knows which settings to alter. To do this, select “Pick” from the “Trace” drop-down menu or press  in the ribbon. Now “Settings” can be selected from the “Dispersion” drop-down menu or the  icon can be selected in the ribbon. In this “Dispersion Analysis” window, select SPAC as the analysis method and SPAC as the method (as seen in Figure 3.13) and then press Apply and Close. Note that there are other options that can be changed, but the default choices proved to be appropriate for this study. The Rayleigh wave velocity at this site did not surpass 800 meters per second and the 150 Hz cap on frequency proved to capture a broad enough range of waves.

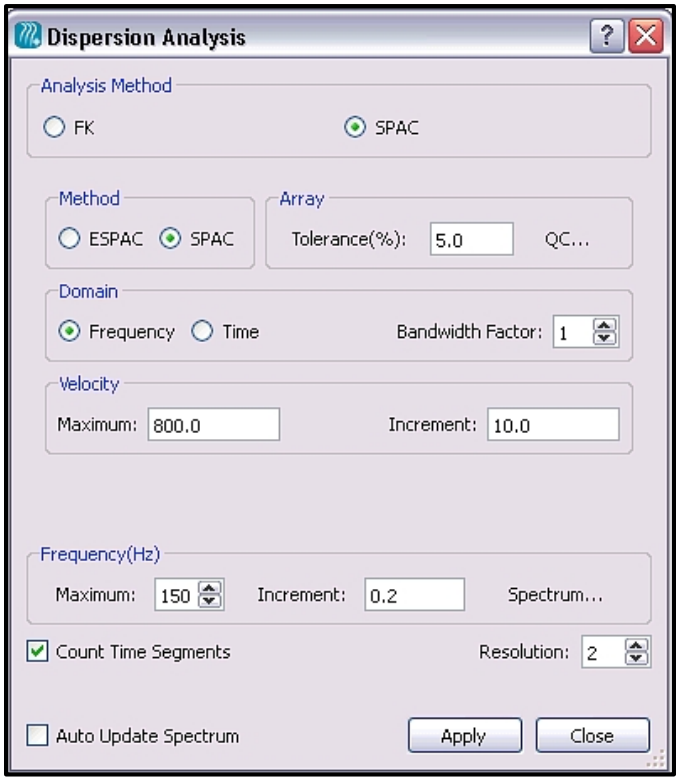


Figure 3.13 Dispersion settings in Geogiga Seismic Pro

At this point, the three recordings can be combined. To do this, select “Combine with Other Records” from the “Dispersion” drop-down menu or select the  icon from the ribbon. In this spectra combination window, highlight each recording in the “Primary” column and select the “Build” option for that column. This should display each recording’s dispersion spectrum – as seen in Figure 3.14. Press Combine and Close to view the combined dispersion spectrum.

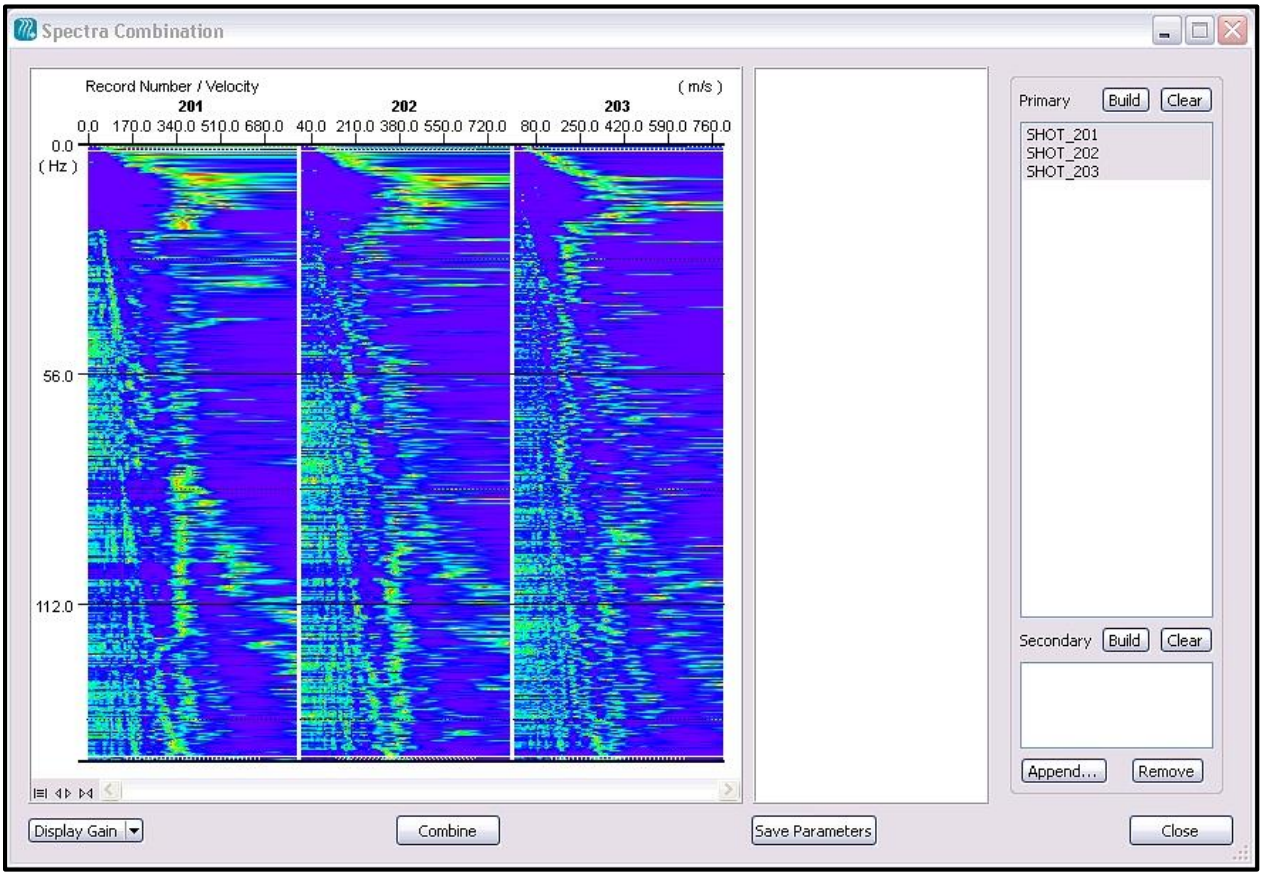


Figure 3.14 Spectra combination in Geogiga Seismic Pro to combine 3 dispersion plots

Next, select points on the dispersion curve by selecting “Picking” from the “Dispersion” drop-down menu or by selecting the  icon in the ribbon. This will bring up a window to select the type of picking; select “Separate” instead of “Continuous” and then close the window.

At this point, the dispersion curve can be created by using the cursor to select points along the spectrum. In normally dispersive situations, the curve should resemble an exponential decay model; pick points that appear to align with this exponential curve. As a rule of thumb for passive surface wave testing, it is better to pick points along the lower bound of this exponential curve instead of the mean value (Louie, 2001). This gives lower, more conservative estimates of the phase velocity and tries to avoid the selection of higher modal responses. For the amount of points to pick, the inversion process will work well as long the points on the dispersion curve exceed the amount of layers trying to be estimated (Rix and Leipski, 1991).

Once these points are picked, the program will automatically create a plot of the Rayleigh wave phase velocity with depth. The depth in this plot is not a direct measure, but rather is calculated from an approximation that assumes that the sampling depth corresponds to half of the wavelength (Rix and Leipski, 1991). An example of the picking and of the corresponding windows can be seen in Figure 3.15. Notice that this curve deviates from the exponential model at higher velocities due to a stiff layer near the ground surface. There is also a blue line in the central window that corresponds to the apparent shear wave velocity values. These values are not found through inversion, but rather through an equation that relates Rayleigh wave velocity to shear wave velocity. This line can be included by clicking the “Curves” drop-down menu, selecting “Dispersion Curve Display”, and then checking the box for Apparent Vs.

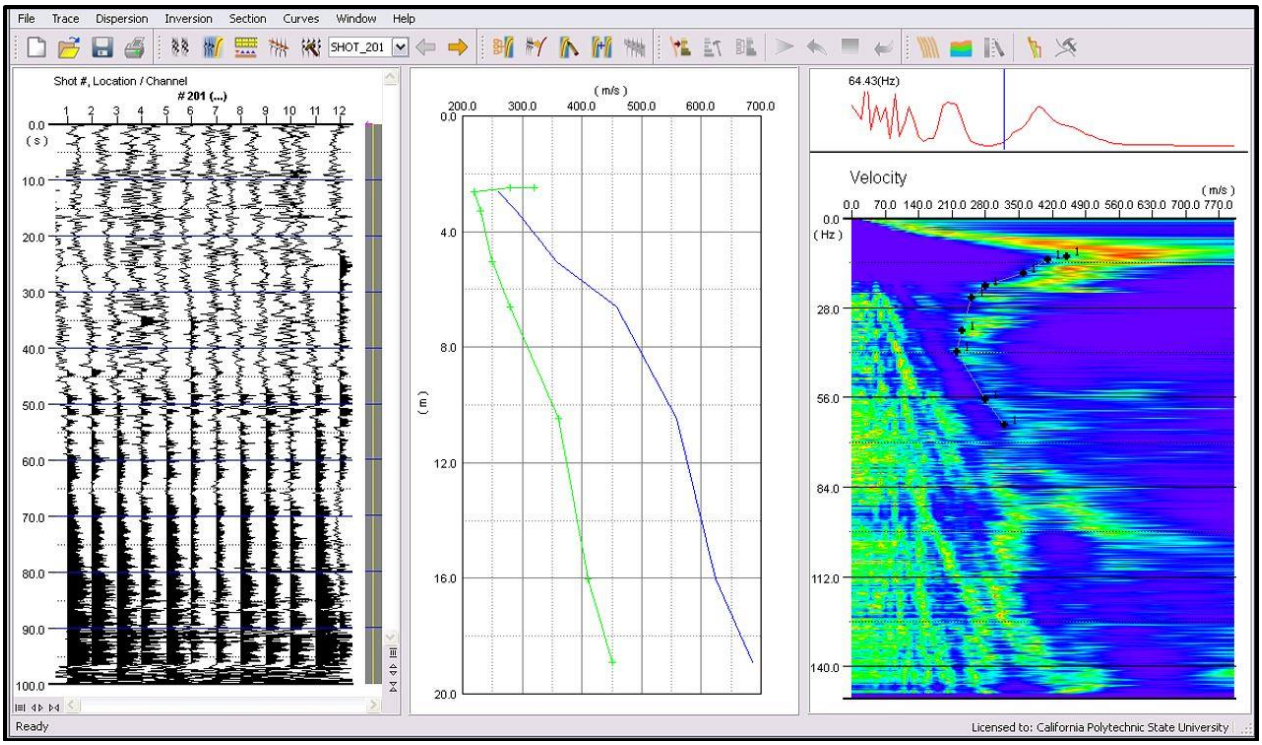


Figure 3.15 Trace, H/V, and dispersion curve plots in Geogiga Seismic Pro

The final step is to save the images of the plots. This can be done by selecting "Save Image" from the "File" drop-down menu. The "View List" column shows which plots can be saved. Highlight the desired plot, name the file, and press Save. This window can be seen in Figure 3.16. Note that, in order to save the image in JPG format, the "Files of Type" drop-down menu needs to be changed to "All files(*)" and the file name should end with ".jpg".

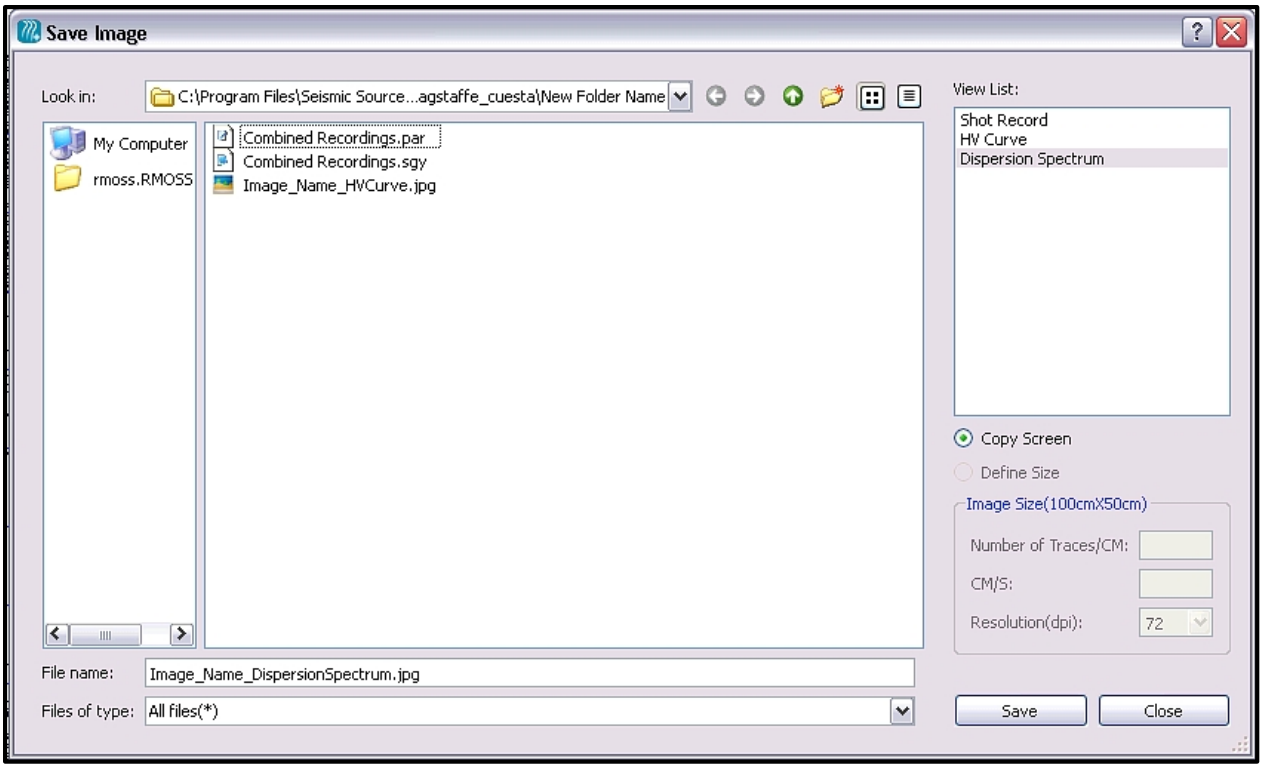


Figure 3.16 Saving an image in Geogiga Seismic Pro

3.5 The Inversion Process

After picking points on the combined dispersion spectrum, the shear wave velocity profile can be created by carrying out an inversion process. Geogiga Seismic Pro has preset algorithms that will carry out this process, but the solution to the inversion problem is not unique (Foti et al., 2014). Many different dispersion curves may correspond to a given shear wave velocity profile. A priori information such as nearby borings, past surveys, or geology maps will create good expectations for depths of layers and shear wave velocity values which can help to mitigate the nonunique solution.

To start the process, follow the steps outlined in the Signal Processing section (above) so that the points on the combined dispersion spectrum are selected. Through forward modeling, it will approximate a shear wave velocity profile that models the selected dispersion curve. The process tries to match the points on the dispersion curve to an estimated model through a least

squares technique. Therefore, it is important to pick a dispersion curve that is representative. An example of this can be seen in Figure 3.17, where the same dispersion spectrum is shown but Figure 3.17 (a) shows points that were poorly selected and Figure 3.17 (b) shows points that are more reasonable for this site. Reasonable curves are typically ones that are normally dispersive since many sites fit into this category, but sites with high-velocity layers overlying lower-velocity layers should reflect this within the dispersion curve. For example, the dispersion curve in Figure 3.17(b) shows a higher-velocity layer at shallow depths. For a discussion on normally dispersive behavior and how dispersion curves vary outside of this behavior, please see Section 2.1.3. The main concern is that appropriate curve should be based on the geology of the site, and inconsistent curves can be avoided if these expectations are known.

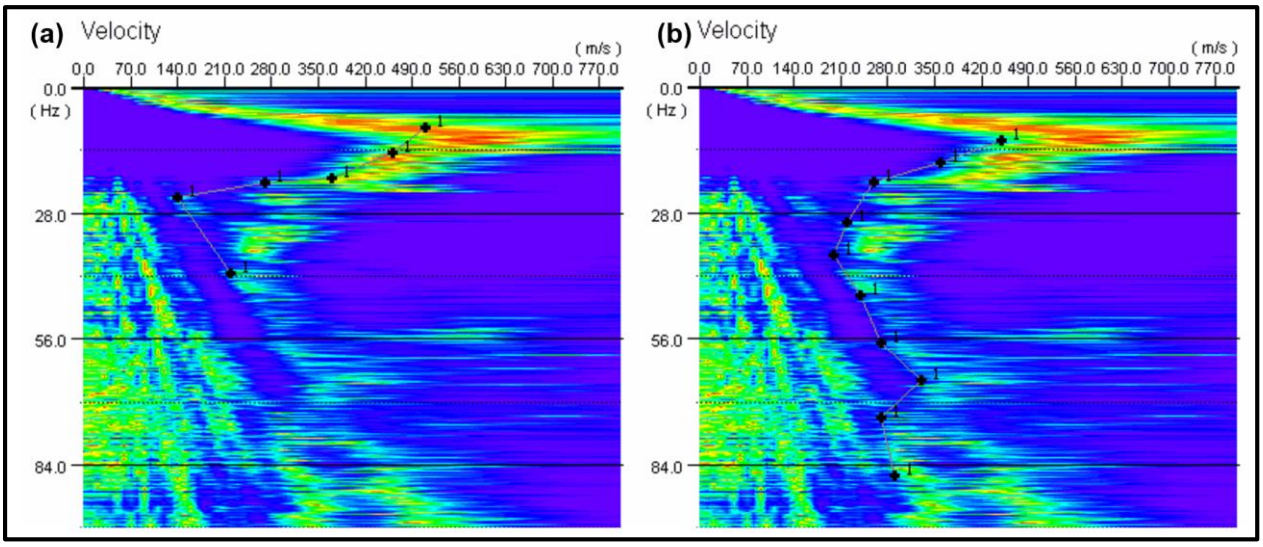


Figure 3.17 (a) Poor dispersion curve picking, and (b) appropriate dispersion curve picking

Once a reasonable curve is picked, click on "Select" from the "Inversion" drop-down menu to indicate that those points are the ones to use for the inversion. Then click on "Autobuild Initial Model" from the same "Inversion" drop-down menu. This will bring up three windows – one showing the initial shear wave velocity profile from the inversion, one showing the dispersion curve that the process is trying to model, and one showing the error between the

selected dispersion curve and the modeled dispersion curve. The default settings only allow the shear wave velocities to vary in the velocity direction. To allow the velocity and the depth to vary in the inversion iterations, click on the "Inversion" drop-down menu and select "Parameters." Then check the area for "Velocity and Depth" as seen in Figure 3.18.

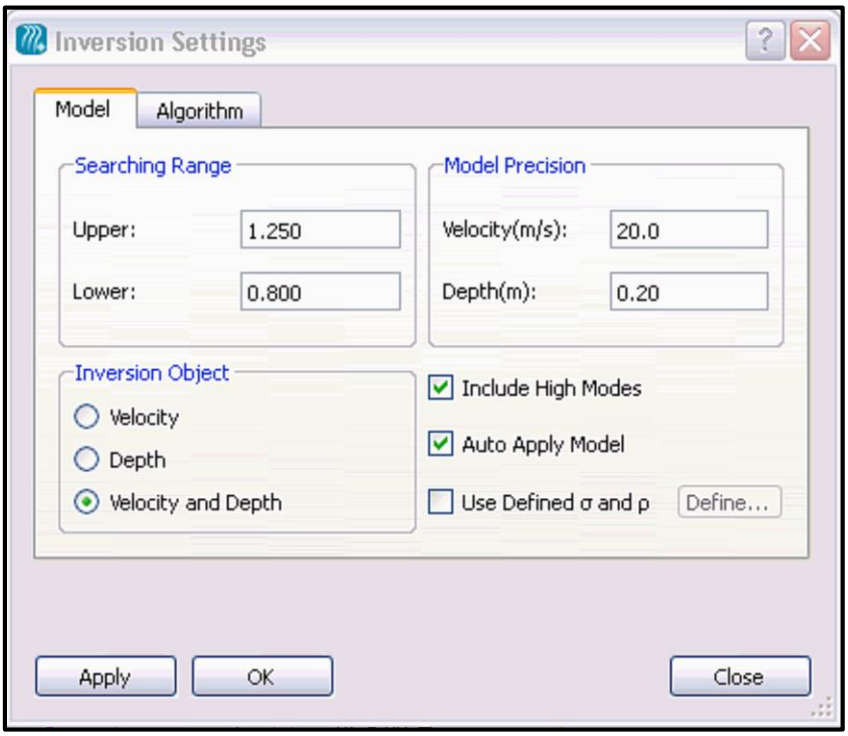


Figure 3.18 Allowing the inversion process to vary by velocity and depth

Since there is no unique solution to this inversion problem, creating the "correct" model is subjective and cannot fully be known unless confirmed by other methods. Because of this uncertainty, the way to create a model that has the most likelihood of being representative of the site is to minimize the error between the observed and the modeled dispersion curve, and to use good judgment about the values the inversion process generates. If no information about the site is known, an assumption that the site is normally dispersive can be a good preliminary guide since geologic processes tend to create this type of stratigraphy. Knowledge of what velocity values to expect is always preferred though and can give bounds to the anticipated profile. Tables

on shear wave velocity values based on geologic units, surrounding topography, or nearby surveys can help to give these bounds. An example of one of these tables can be seen in Table 3.1 which shows the Pacific Earthquake Engineering Research Center's (PEER) description of the shear wave velocity characteristics of geologic units in California from Wills and Clahan (2006).

While being observant of the velocity values, the error between the observed/selected dispersion curve and the modeled dispersion curve can be lessened through Geogiga Seismic Pro by selecting "Start" from the "Inversion" drop-down menu or by pressing  in the ribbon. This will run the inversion process again with the initial values being those created from the "Autobuild Initial Model." The additional iterations should reduce the error between the models and produce a new shear wave velocity profile in red. Note that layers can be removed from or added to the model by holding down Ctrl and selecting a layer.

Geologic Unit	Geologic Description	No. of Profiles	Mean $V_{s,30}$	Std Dev	$V_{s,30}$ from Mean of ln	Std Dev of ln	Mean of ln of $V_{s,30}$
Qi	Intertidal Mud, including mud around the San Francisco Bay and similar mud in the Sacramento/San Joaquin delta and in Humboldt Bay	20	160	39	155	0.243	5.046
af/qi	Artificial fill over intertidal mud around San Francisco Bay	44	217	94	202	0.357	5.310
Qal, fine	Quaternary (Holocene) alluvium in areas where it is known to be predominantly fine	13	236	55	229	0.238	5.437
Qal, deep	Quaternary (Holocene) alluvium (Holocene and Pleistocene) is more than 30 meters thick; generally much more in deep basins	161	280	74	271	0.250	5.604
Qal, deep Imperial Valley	Quaternary (Holocene) alluvium in the Imperial Valley, except sites in the northern Coachella Valley adjacent to the mountain front	53	209	31	207	0.135	5.335
Qal, deep LA Basin	Quaternary (Holocene) alluvium in the Los Angeles basin, except sites adjacent to the mountain fronts	64	281	85	270	0.275	5.599
Qal, thin	Quaternary (Holocene) alluvium in narrow valleys, small basins, and adjacent to the edges of basins where the alluvium would be expected to be underlain by contrasting material within 30 meters	65	349	89	338	0.244	5.825
Qal, thin West LA	Quaternary (Holocene) alluvium in part of west Los Angeles where the Holocene alluvium is known to be thin, and is underlain by Pleistocene alluvium	41	297	45	294	0.150	5.684
Qal, coarse	Quaternary (Holocene) alluvium near fronts of high, steep mountain ranges and in major channels where the alluvium is expected to be coarse	18	354	82	345	0.223	5.845
Qoa	Quaternary (Pleistocene) alluvium	132	387	142	370	0.273	5.916
Qs	Quaternary (Pleistocene) sand deposits, such as the Merritt Sand in the Oakland area	15	302	46	297	0.171	5.697
QT	Quaternary to Tertiary (Pleistocene-Pliocene) alluvial deposits such as the Saugus Formation of southern California, Paso Robles Formation of central coast ranges, and the Santa Clara Formation of the Bay Area	18	455	150	438	0.266	6.083
Tsh	Tertiary (mostly Miocene, Oligocene, and Eocene) shale and siltstone units such as the Repetto, Fernando, Puente, and Modelo Formations of the Los Angeles area	55	390	112	376	0.272	5.930
Tss	Tertiary (mostly Miocene, Oligocene, and Eocene) sandstone units such as the Topanga Formation in the Los Angeles area and the Butano sandstone in the San Francisco Bay area	24	515	215	477	0.386	6.169
Tv	Tertiary volcanic units including the Conejo Volcanics in the Santa Monica Mountains and the Leona Rhyolite in the East Bay Hills	3	609	155	597	0.240	6.392
Kss	Cretaceous sandstone of the Great Valley Sequence in the central Coast Ranges	6	566	199	539	0.332	6.291
Serpentine	Serpentine, generally considered part of the Franciscan complex	6	653	137	641	0.204	6.464
KJf	Franciscan complex rock, including mélange, shale, chert, and greenstone	32	782	359	712	0.432	6.569
xtaline	Crystalline rocks, including Cretaceous granitic rocks, Jurassic metamorphic rocks, schist, and Precambrian gneiss	28	748	430	660	0.489	6.493

Table 3.1 Shear wave velocity characteristics of geologic units in CA from Wills and Clahan

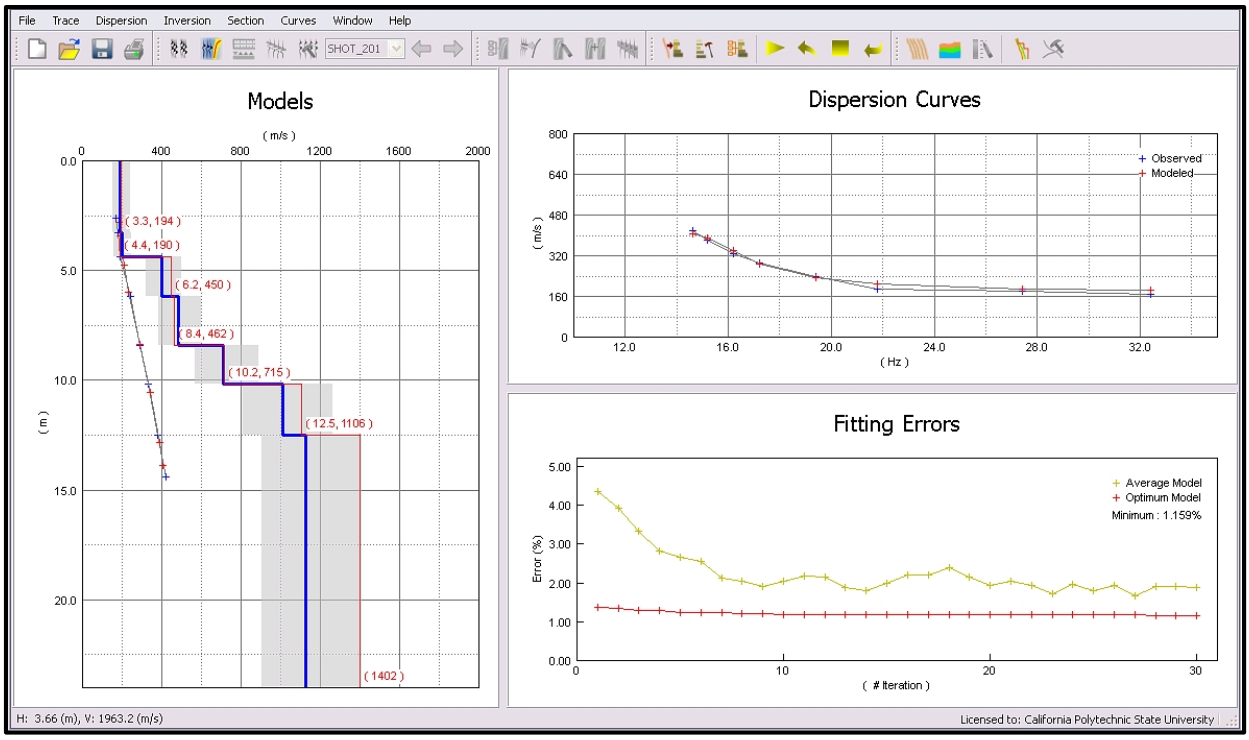


Figure 3.19 The shear wave velocity profile, dispersion curve, and error iteration windows

This process can be repeated by pressing “Reset” in the “Inversion” drop-down menu or by pressing  in the ribbon. This changes the initial model to be the one just created (in red), and then pressing “Start” from the “Inversion” drop-down menu will run more iterations. This inversion process should continue to lower the error and make the dispersion curves more similar. However, a representative profile is the goal of this process, not a low error rate, and, although these can be related, too many iterations tend to produce obscure results. This is when expectations of the geology can help decide when the model shows a reasonable profile along with a suitably low error rate. Since it is hard to know when to stop without actually running too many iterations, trial and error is often the most helpful resource. Through this trial and error, a calculation of the uncertainty in the inversion results can help constrain the data (could calculate the median value and iterate to this value). If the profile has gone through too many iterations,

“Autobuild Initial Model” and then the “Reset” button can be selected in order to start from the beginning.

To keep the shear wave velocity models consistent, this study created profiles using the same procedure for each location. This procedure is site specific and was generated by noting how the velocity profiles changed with several iterations at multiple locations. Three layers were used unless the dispersion curve showed evidence of an inverted layer. Iterations were stopped when the error dropped below 2% or if the velocity values started to vary significantly from adjacent surveys. If the error after the first round of iterations was below 2%, then one more iteration was complete. Since the recordings for this study overlapped, this procedure was created in an attempt to make neighboring profiles similar to one another, and to keep velocity values reasonable based on the known geology.

Chapter 4: Results

4.1 SPAC Results

Each survey was taken at 5-meter spacing following the procedures outlined in Chapter 3. The approximate testing locations can be seen in Figure 3.2, and the results of the inversion procedure are summarized in Figure 4.1. Note that the vertical scale on this figure has been exaggerated so that the velocity values can be seen more easily. All of the graphs for the individual surveys can be seen in the Appendix.

Three depths were analyzed based on the amount of available data. The average sample depth was from 3.2 to 17.3 meters from the ground surface, so depths of 5, 10, and 15 meters from the ground surface were analyzed. A semi-variogram was also created for data at 20 meters depth, but the small amount of data at this depth made the plot too erratic. Note that, since there is an elevation change at the site, all depth values in this study are in reference to the ground surface in the valley (not on the hill).

This study used the time weighted average velocity since this is what is used in practice (and required by the NEHRP site classification guidelines). This allows a column of geologic material to be analyzed instead of point values. It is estimated by dividing the depth by the amount of time it took a wave to propagate through each layer within that depth. The equation can be written as;

$$\bar{V} = \frac{d_T}{\sum_{i=1}^{n} \frac{d_i}{V_i}}$$

where $\bar{V}$ is the time weighted average velocity of the column, d_T is the total depth of interest, d_i is the thickness of each layer within d_T, V_i is the velocity of each layer, and n

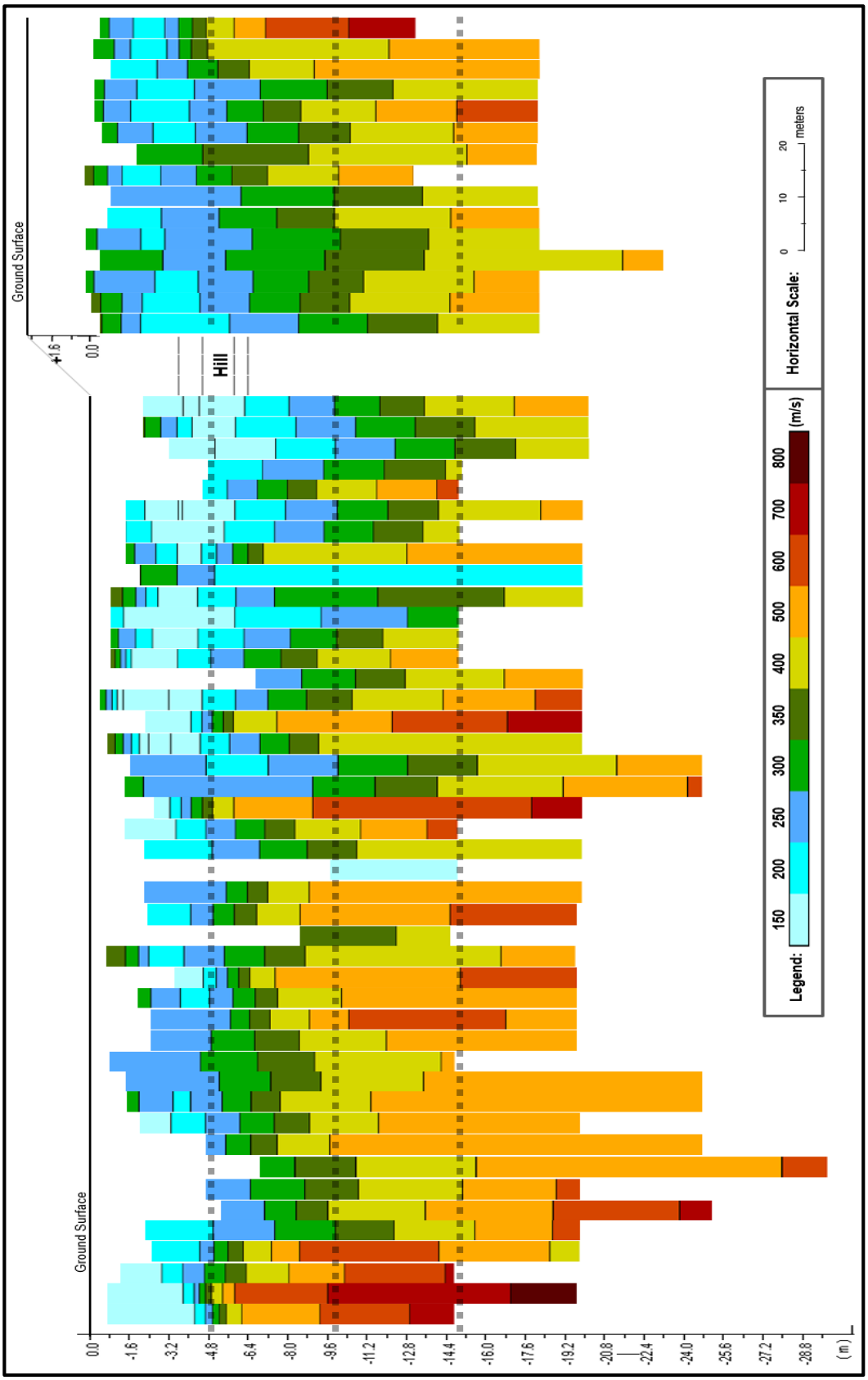


Figure 4.1 Shear wave velocity profile for each survey with dotted lines at analyzed depths

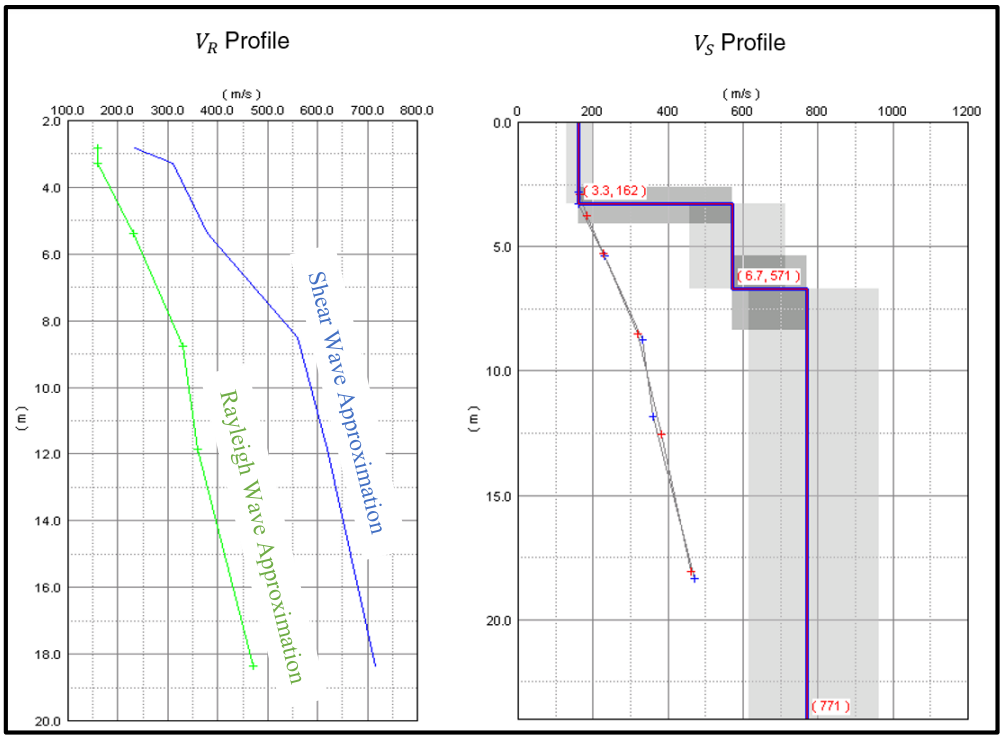


Figure 4.2 Example of Rayleigh wave and shear wave velocity profiles from Geogiga (2012)

represents the total number of layers. This equation is easy to calculate when using shear wave velocity profiles because of their "stepped" nature, but the Rayleigh wave velocity profile is estimated via the dispersion curve (from Rix and Leipski, 1991, half-wavelength approximation) which does not have constant, layered velocity values. A comparison can be seen if Figure 4.2. In order to analyze the Rayleigh wave velocity profiles, the same concept was applied. Each segment was treated as a layer and average velocity values were used for each layer. This weights the velocity based on the average time it took waves to propagate through each segment. The results for both the Rayleigh wave and shear wave velocity values can be seen in the Appendix.

The time weighted average Rayleigh wave velocities ($\bar{V}_R$) at 5 meters, 10 meters, and 15 meters depth can be seen in Figure 4.3. Similarly, the time weighted average shear wave velocities ($\bar{V}_S$) at these depths can be seen in Figure 4.4. The offsets in the figure are related to

the reference stake (see Figure 3.2) where offsets north of this stake are negative and south of this stake are positive. The gap in the data corresponds to the 3-array break due to the hill seen in Figure 3.2. The higher velocity values past the hill are assumed to be due to the additional overburden stress that compressed the material below.

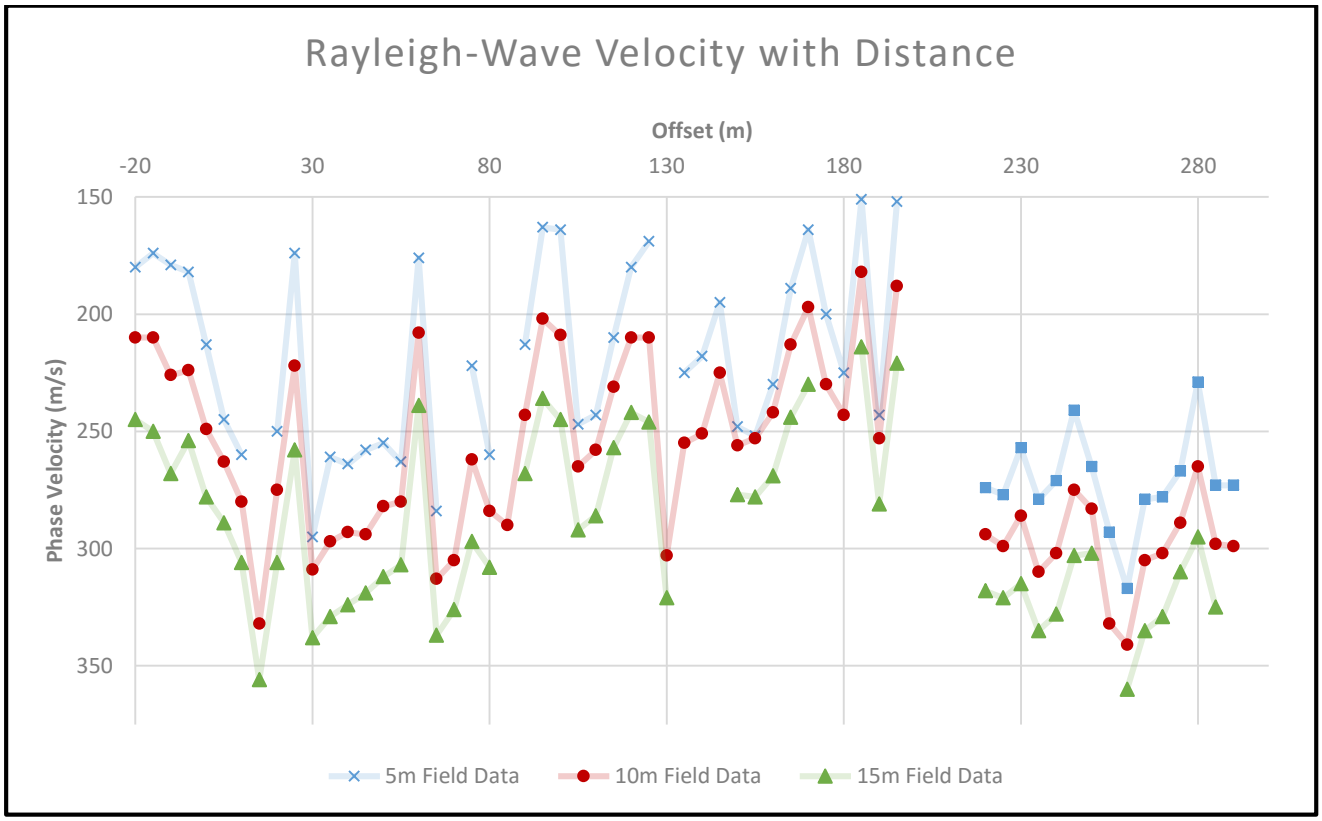


Figure 4.3 Time-averaged Rayleigh wave velocity across the site

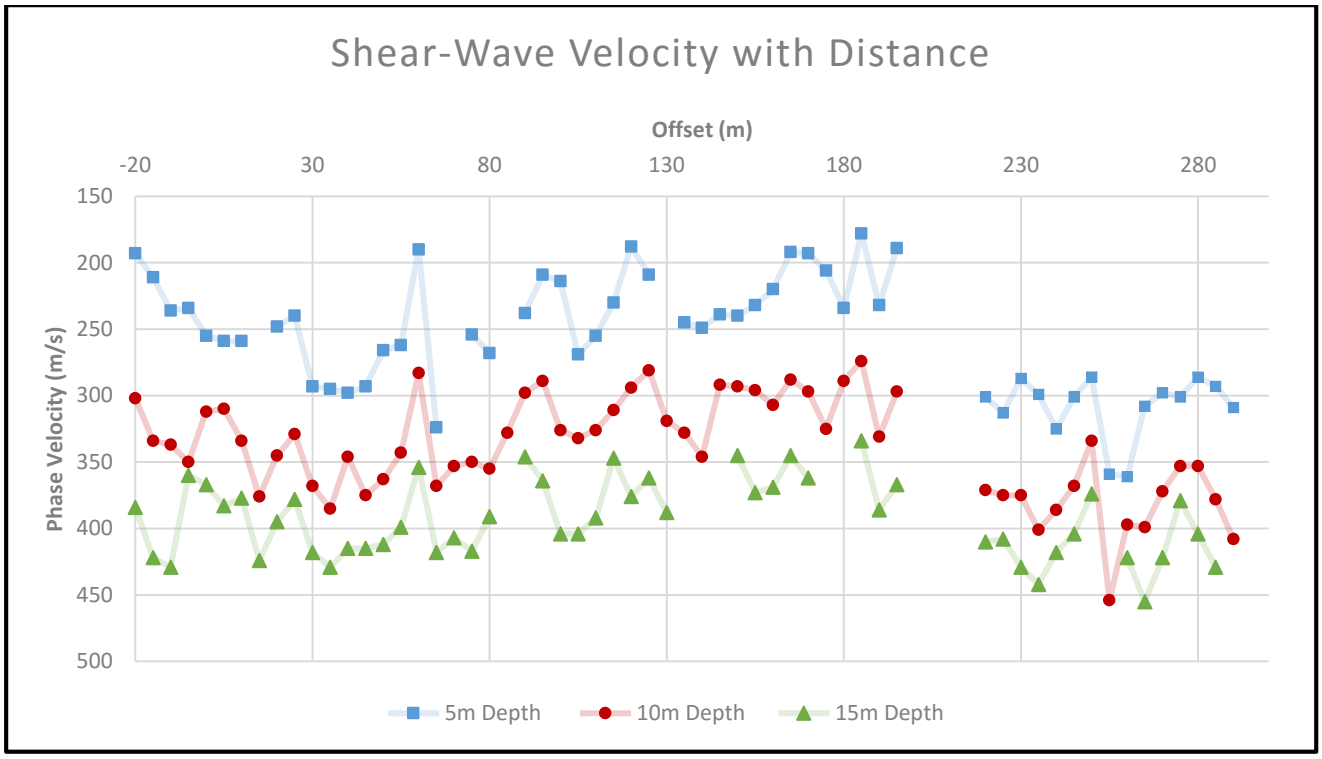


Figure 4.4 Time-averaged shear wave velocity across the site

4.2 Empirical Semi-Variograms and Covariance Functions

The semi-variogram and covariance function using all of the $\bar{V}_R$ data can be seen in Figure 4.5 with the corresponding number of data pairs used for each point. Note that the numerical subscripts refer to the analyzed depth. The $\bar{V}_{R,10}$ plots are the most constrained which appears to be from the additional data and the geologic similarities at this depth. The covariance function, however, flattens at negative values, and this behavior is indicative of a trend/non-stationarity in the data (Oliver and Webster, 2010). In other words, the mean value is not constant throughout the site.

The seemingly obvious source of the apparent trend is from the distinct increase in velocity values after the hill. In attempts to rid the data of this trend, the "hill" data was removed from the set and re-analyzed; these results can be seen in Figure 4.8. Another option for removing this apparent trend is to fit a trendline to the data and then create a semi-variogram

from the residuals (Oliver and Webster, 2015). This tries to model the changing mean value so that the spatial statistics are calculated with a pseudo-constant mean value. A polynomial function and a piecewise linear function were used as the trendlines because these fit closest to the data. The largest R^2 value (while still being parsimonious) was from a fourth-order polynomial function and the piecewise-linear function was split into sections before the hill and after the hill. The trendlines can be seen in Figure 4.6 and Figure 4.7, respectively, and the geostatistical results can be seen in Figure 4.9 and Figure 4.10, respectively.

When looking at the piecewise-linear plot, the data could also be modeled as a simple linear function without the "hill" data (i.e. the first portion of the piecewise-linear function). This models the data in case there was a trend within the valley. These results can be seen in Figure 4.11.

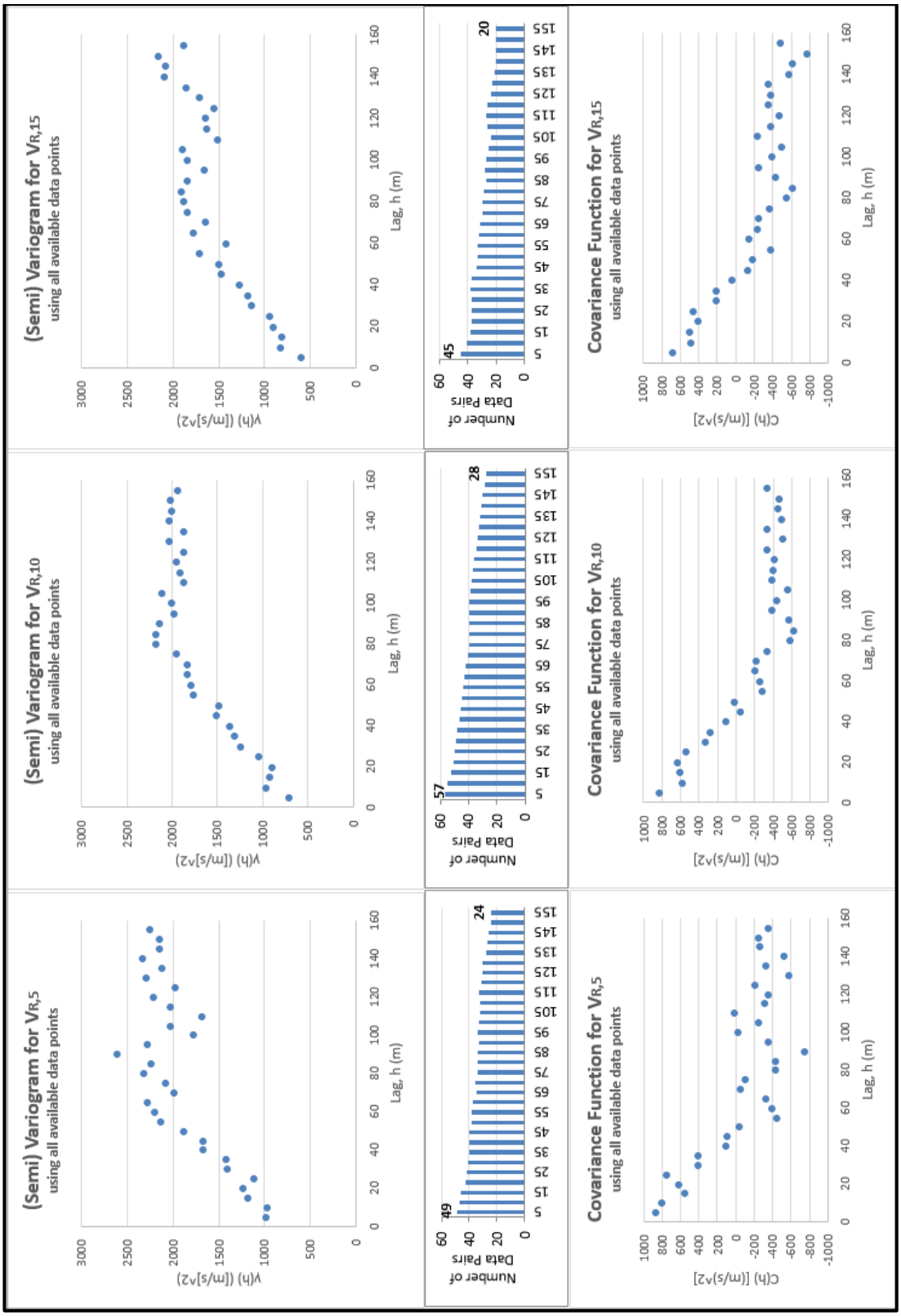


Figure 4.5 Semi-variogram and covariance function for time weighted average Rayleigh wave velocity using all available data points

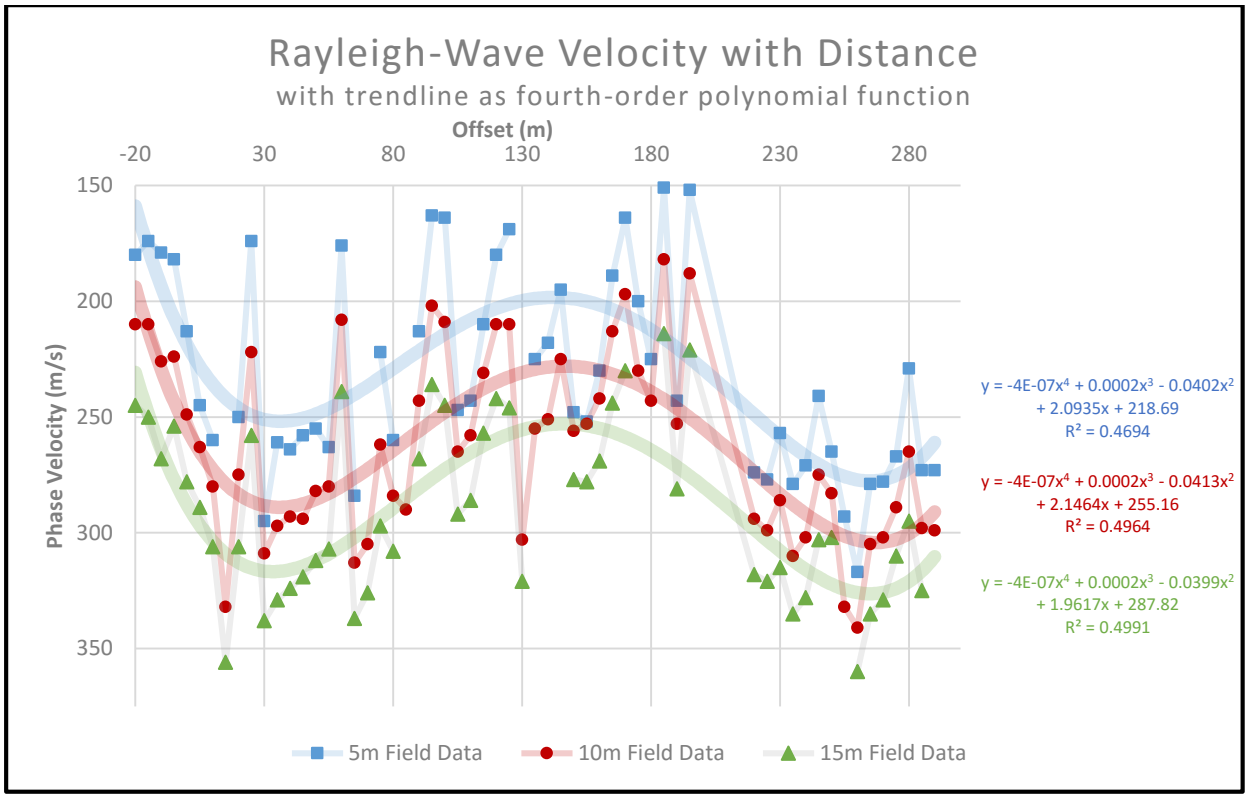


Figure 4.6 Time weighted average Rayleigh wave velocity fit with fourth-order polynomial functions

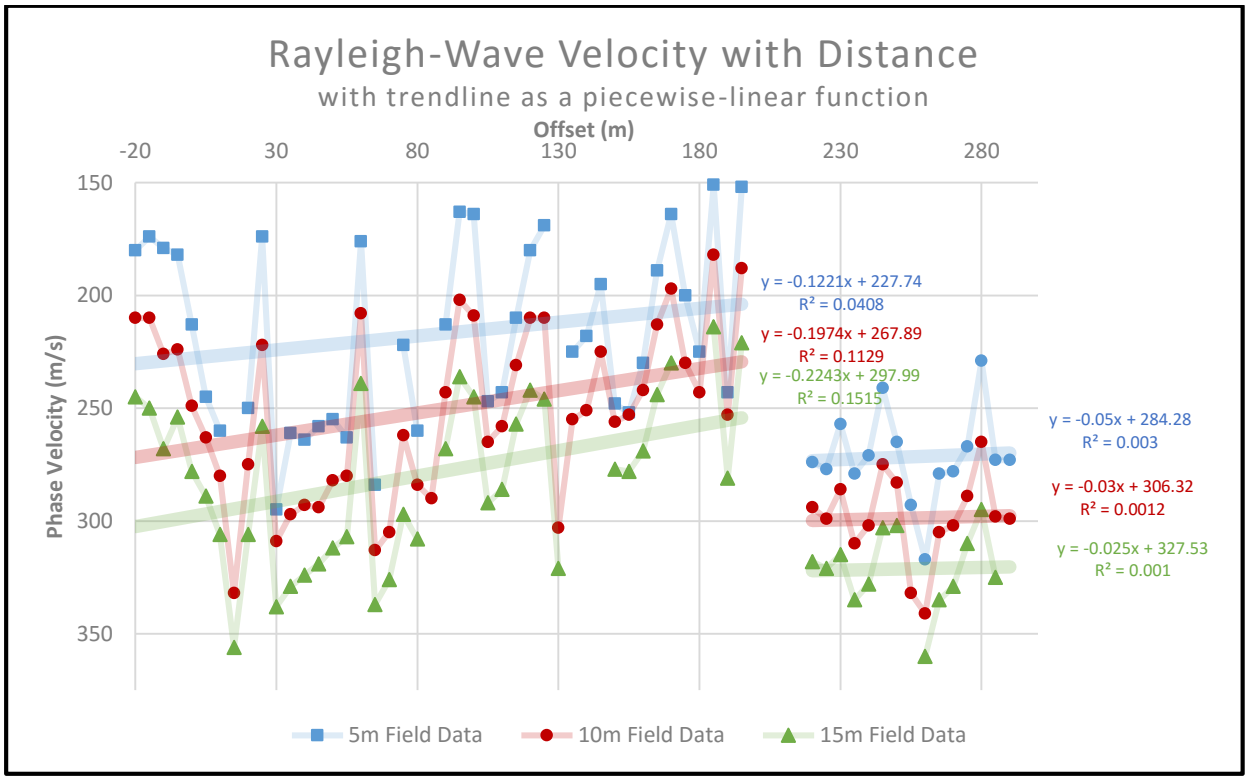


Figure 4.7 Time weighted average Rayleigh wave velocity fit with piecewise-linear functions

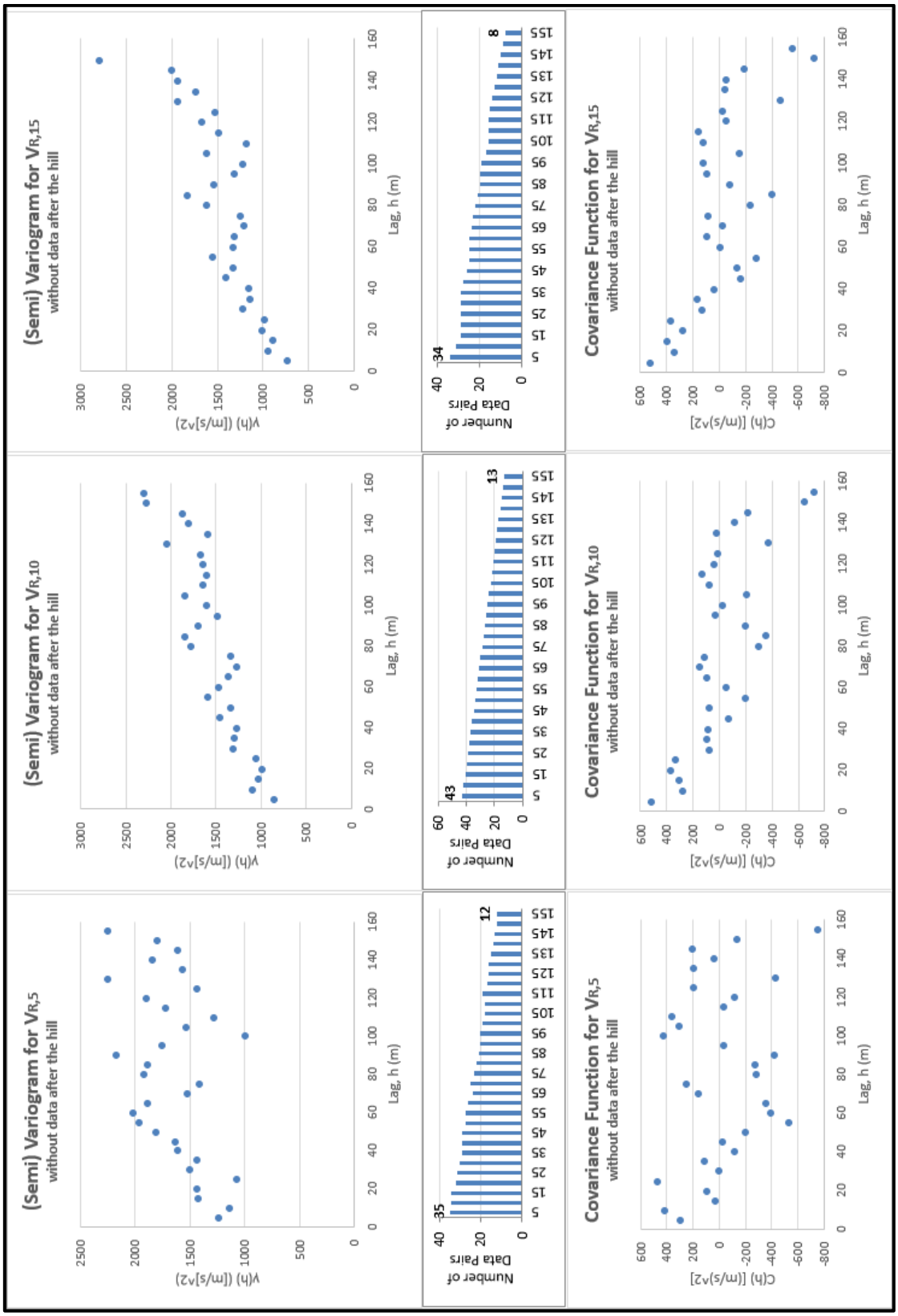


Figure 4.8 Semi-variogram and covariance function for time weighted average Rayleigh wave velocity not including data from after the hill

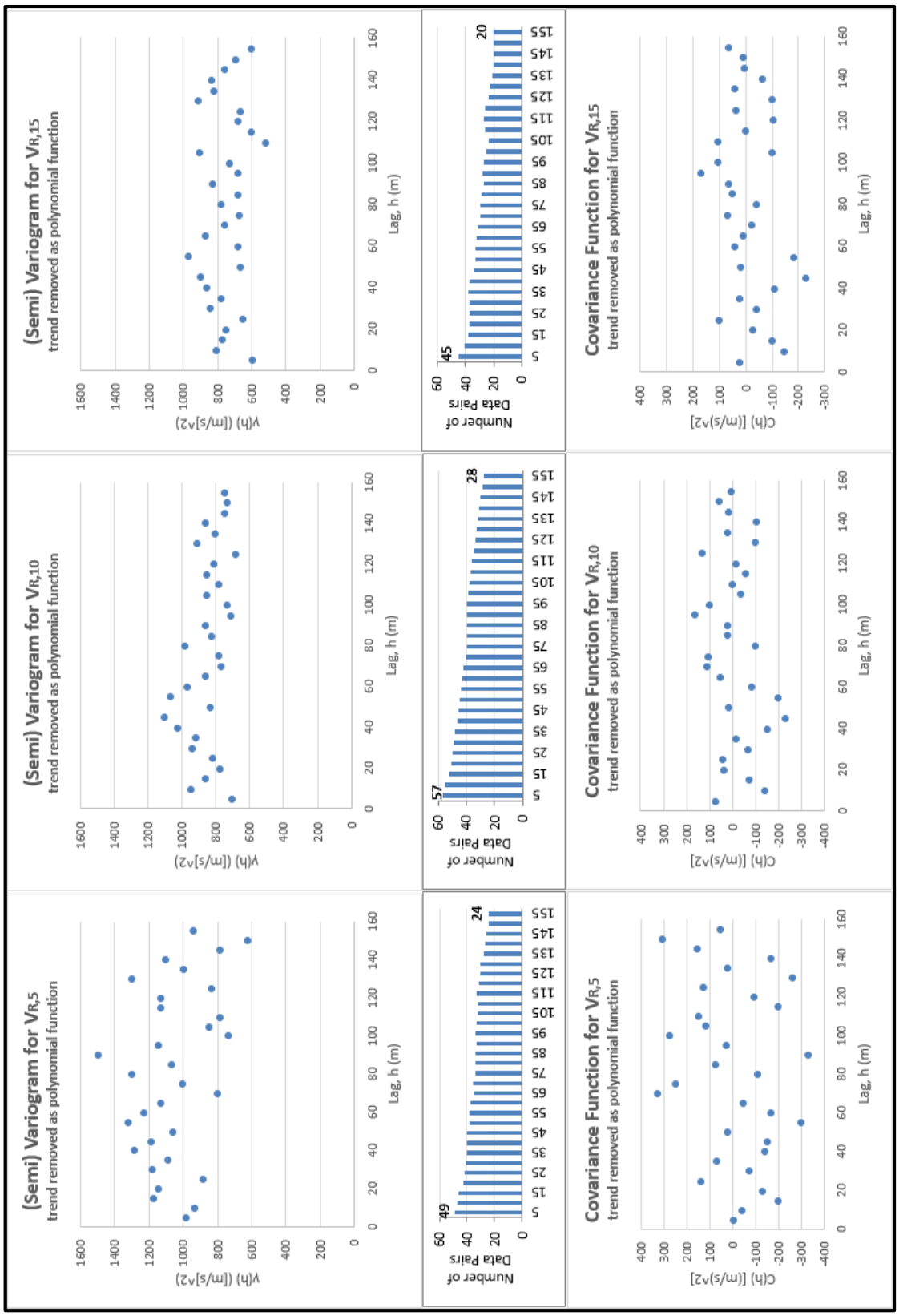


Figure 4.9 Semi-variogram and covariance function for time weighted average Rayleigh wave velocity with trend removed as a fourth-order polynomial function

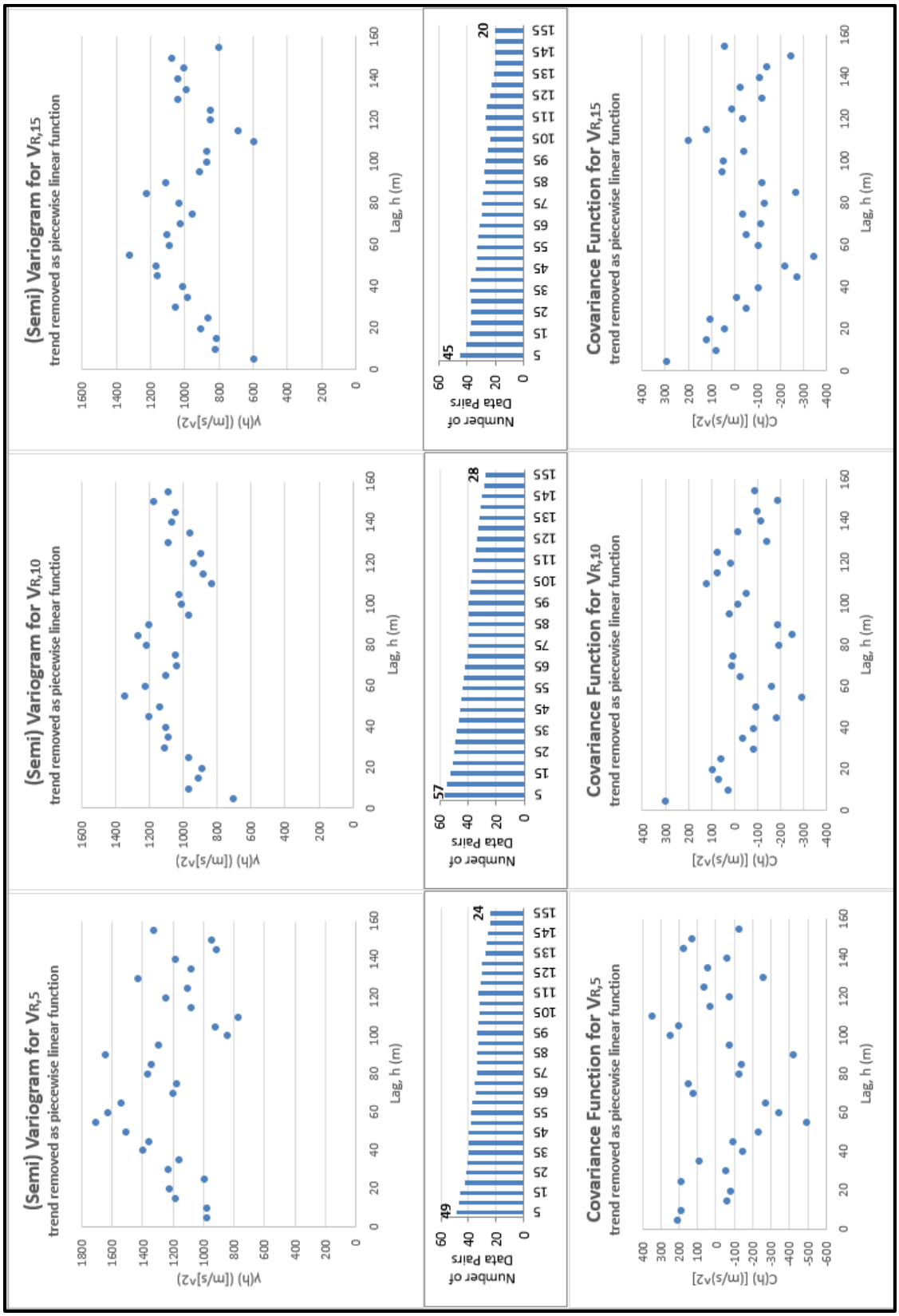


Figure 4.10 Semi-variogram and covariance function for time weighted average Rayleigh wave velocity with trend removed as a piecewise-linear function

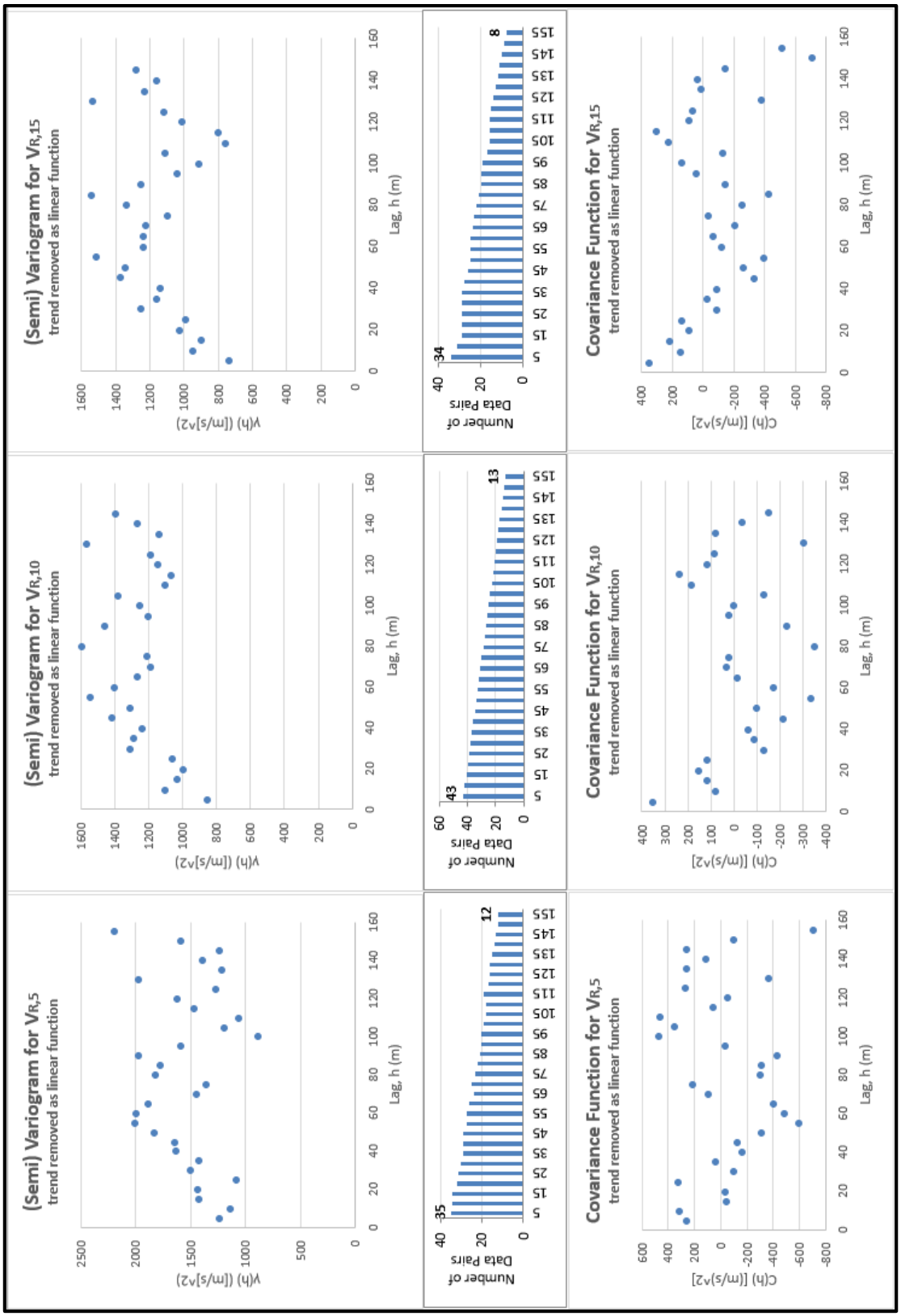


Figure 4.11 Semi-variogram and covariance function for time weighted average Rayleigh wave velocity with trend removed as a linear function (without data after hill)

4.3 Modeling the Semi-Variograms

Notice that the plots become more erratic as the amount of data pairs decreases – this is typical (Oliver and Webster, 2015). Also note that the semi-variograms using the residuals do not appear to constrain the data any better. This can be seen more easily through modeling the data. All of these semi-variograms were modeled with spherical, exponential, and Gaussian models. These three models were chosen because they are the common transition models (models including a sill) and a sill is expected for alluvial material (e.g. Iqbal et al., 2005, and Facciorusso et al., 2010). Oliver and Webster (2015) recommend choosing a semi-variogram model using the coefficient of determination (R^2) and any knowledge of the geologic behavior. The R^2 values for each semi-variogram model and the different remedies for the apparent trend can be seen in Table 4.1. Additionally, the semi-variograms, covariance functions, and correlograms for each of these options can be seen in the Appendix.

		Rayleigh wave Velocity Data			Shear wave Velocity Data		
Data Used	Model	5m Depth	10m Depth	15m Depth	5m Depth	10m Depth	15m Depth
1. All Available Data	Spherical	**83%**	**95%**	**86%**	**94%**	**91%**	**72%**
	Exponential	**79%**	**91%**	**86%**	88%	87%	70%
	Gaussian	**84%**	**95%**	**86%**	**94%**	**92%**	**73%**
2. Data after Hill Removed	Spherical	**36%**	76%	**68%**	**68%**	78%	50%
	Exponential	**31%**	76%	**68%**	**65%**	76%	49%
	Gaussian	**35%**	74%	**68%**	**69%**	79%	54%
3. Trend Removed as Linear Function without Data after Hill	Spherical	**23%**	**57%**	**48%**	**34%**	**27%**	3%
	Exponential	**19%**	**55%**	**45%**	**30%**	**24%**	3%
	Gaussian	**22%**	**57%**	**48%**	**33%**	**27%**	3%
4. Trend Removed as Piecewise Linear Function	Spherical	0.8%	**24%**	**21%**	39%	**8%**	0.4%
	Exponential	**8%**	**24%**	**21%**	42%	**8%**	0.6%
	Gaussian	0.5%	**23%**	**20%**	41%	**7%**	0.8%
5. Trend Removed as 4th Order Polynomial	Spherical	7%	15%	0.3%	16%	7%	4%
	Exponential	4%	14%	0.3%	15%	7%	4%
	Gaussian	6%	6%	0.6%	2%	7%	4%
Key: **Sill is visible within the data**; Data resembles the nugget effect model							

Table 4.1 Coefficient of determination values for each data type and semi-variogram model

Data Used		Rayleigh wave Velocity Data			Shear wave Velocity Data		
		5m Depth	10m Depth	15m Depth	5m Depth	10m Depth	15m Depth
1. All Available Data	Spherical	**81**	**92**	**89**	**97**	**131**	**142**
	Gaussian	**50**	**75**	**60**	**70**	**100**	**100**
2. Data after Hill Removed	Spherical	**63**	129	**87**	**87**	270	399
	Gaussian	**50**	90	**50**	**60**	-	-
3. Trend Removed as Linear Function without Data after Hill	Spherical	**57**	**58**	**48**	**60**	**59**	1668
	Gaussian	**30**	**45**	**30**	**40**	**55**	-
4. Trend Removed as Piecewise Linear Function	Spherical	-	**38**	**34**	170	**83**	-
	Gaussian	-	**25**	**25**	**55**	**50**	-
5. Trend Removed as 4th Order Polynomial	Spherical	-	-	-	-	-	-
	Gaussian	-	-	-	-	-	-
Key: **Sill is visible within the data**; Data resembles the nugget effect model							

Table 4.2 Semi-variogram ranges from the spherical and Gaussian models for each data type

In Table 4.1, notice that the exponential model gives a worse fit than the spherical or Gaussian models in almost all the cases, whereas the spherical and Gaussian models are nearly identical in terms of R^2 values. Because of this, the ranges (distance to the sill) were found for the spherical and Gaussian models only. These ranges can be seen in Table 4.2. The range for the spherical models was found using open-source, MATLAB code (Schwanghart, 2010), but this code could not accurately identify the range for Gaussian models. Because of this, the Gaussian models' ranges were estimated manually by taking 95% of the sill value (the definition of the range for models that reach a sill asymptotically).

The results show that all of the Gaussian ranges are shorter than their spherical counterparts. This is most likely due to the fact that the Gaussian models approach the sill at similar lag distances as the spherical models. While spherical models define their range by when

the model reaches the sill, the Gaussian model's range is defined by the 95% of that value and so it will be inherently lower. An example of this can be seen in Figure 4.12.

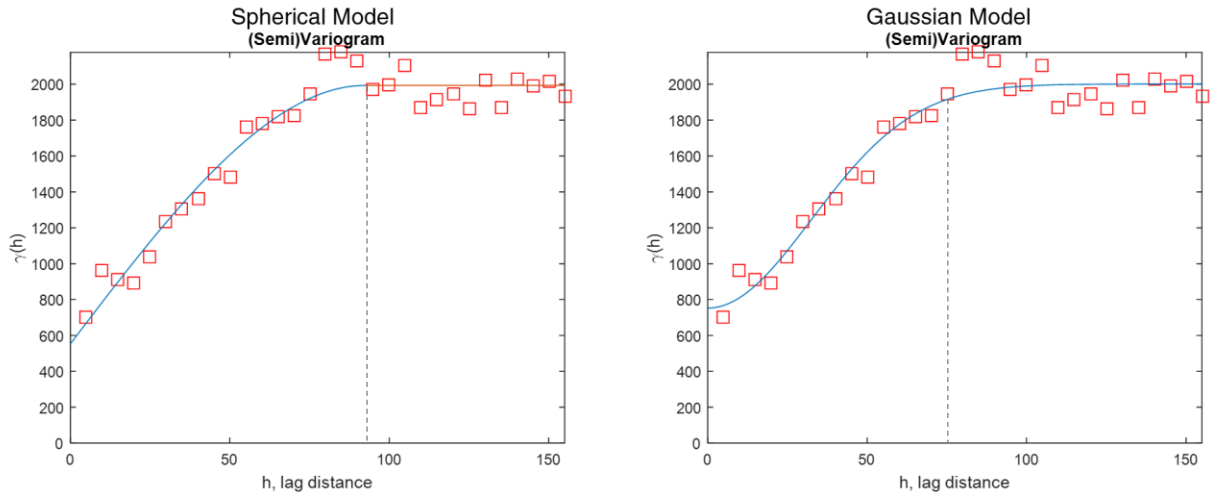


Figure 4.12 Semi-variogram from Rayleigh wave data at 10m depth using all available data and being modeled with a spherical model (left) and a Gaussian model (right)

Note that values that are not bolded in Table 4.1 and Table 4.2 have ranges outside the boundaries of the data set. This implies that a transitional model is not appropriate; a linear model would better represent the data in these cases. However, a linear model is not expected for this data set, has no range, and could be indicative of another trend (Oliver and Webster, 2015). Although, the data in these situations appear to be so erratic that a line is the only model that sufficiently lowers the R^2 value. As a result, these unbolded values should not be considered reliable.

To help give an idea of the shape of these semi-variograms and how well they fit the model, an overview (modeled with spherical models) can be seen in Figure 4.13 where the row numbers correspond to the row numbers in Table 4.1 and Table 4.2. These figures are meant to give a broad sense of how the semi-variograms vary, but, for further inspection, these figures can be seen in detail in the Appendix.

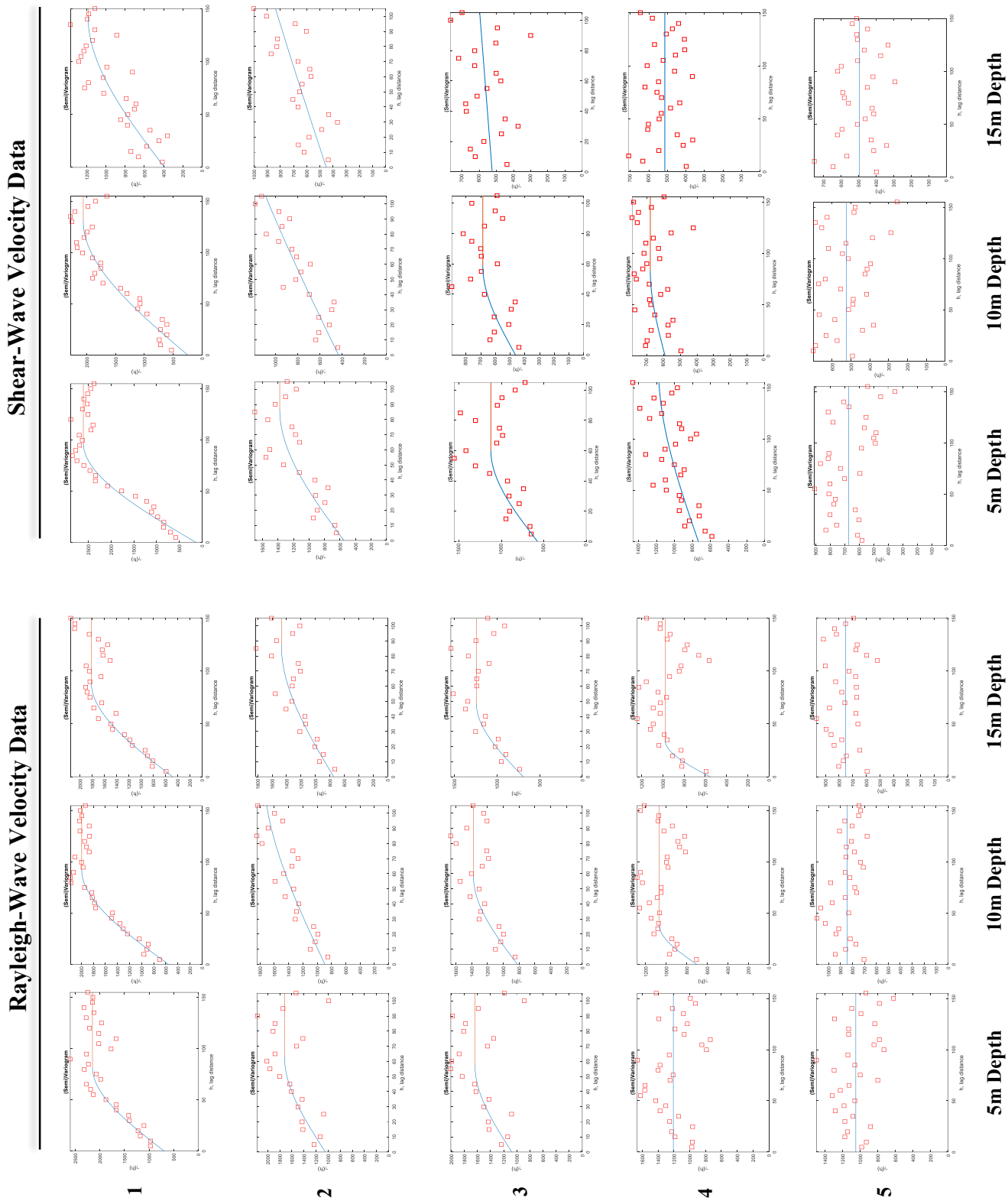


Figure 4.13 Overview of semi-variogram plots fitted with spherical models

4.4 Estimating the Ranges

Given all of this, an applicable range value should be chosen based on all of these factors – the R^2 values, the semi-variogram shape, and the spread of the range values. The only data whose models explain over half of the variation ($R^2 > 50\%$) are those with all the data included, those with the "hill" data removed, and those with the trend removed as a linear function after the "hill" data was removed. A summary of these range values that have R^2 values greater than 50% and whose sills are visible within the data can be seen in Table 4.3.

Data Used		Rayleigh wave Velocity Data			Shear wave Velocity Data		
		5m Depth	10m Depth	15m Depth	5m Depth	10m Depth	15m Depth
All Available Data	Spherical	81 (83%)	92 (95%)	89 (86%)	97 (94%)	131 (91%)	142 (72%)
	Gaussian	50 (84%)	75 (95%)	60 (86%)	70 (94%)	100 (92%)	100 (73%)
Data after Hill Removed	Spherical			87 (68%)	87 (68%)		
	Gaussian			50 (68%)	60 (69%)		
Trend Removed as Linear Function without Data after Hill	Spherical		58 (57%)				
	Gaussian		45 (57%)				

Key: Range (R^2 value); Sill not visible within the data or R^2 value is lower than 50%

Table 4.3 Summary of range values whose model has R^2 values greater than 50% and whose sill is visible within the data

It is hard to trust the values from the last two data portions because none of the adjacent depths are well modeled and the conversion to shear wave velocity is also not well modeled (or the value was estimated from poorly-modeled Rayleigh wave velocity). And, in regards to removing the apparent trend, these two data portions also give negative covariance functions, so they prove to be no more useful than the initial model. Because of this, the only mathmatically reliable data is that which includes all the available data; the R^2 values are high for both Rayleigh and shear wave velocity, the semi-variograms are well constrained, and all depths appear to be well modeled. The problem remains that second-order stationarity is violated (i.e.

that there is not a constant mean value across the site). This indicates that the covariance functions and correlograms are not applicable, however Matheron's intrinsic hypothesis (see Section *2.2.3 Spatial Continuity Models*) can be used which allows the semi-variogram values to be applicable even though the covariance function and correlogram are not (Schekhar, 2008).

Using all the data, the ranges are similar at 10 and 15 meter depths but smaller at 5 meters depth. This seems justified since the near-surface material tend to have more differences than deeper material (from events like seasonal influences or utility installations or differing water tables). Assuming that the range values for 10 and 15 meters depth are similar enough that averaged values accurately represent these depths, the results can be seen in Table 4.4.

		Rayleigh wave Velocity Data		Shear wave Velocity Data	
Data Used	Model	5m Depth	10m and 15m Depth	5m Depth	10m and 15m Depth
All Available Data	Spherical	**81**	**90.5**	**97**	**136.5**
	Gaussian	**50**	**67.5**	**70**	**100**

Table 4.4 Range values using all available data and averaging the values from 10 and 15m depth

Since the R^2 values are nearly identical between the spherical and Gaussian models, the smaller ranges of the Gaussian models will be used for conservatism. For Rayleigh waves, this gives a correlation range of 50 meters for a 5-meter-deep column of alluvial soil and of approximately 65 meters for a 10- to 15-meter-deep column of alluvial soil. For shear waves, this gives a correlation range of 70 meters for a 5-meter-deep column of alluvial soil and of 100 meters for a 10- to 15-meter-deep column of alluvial soil.

Boore (2004) gives a correlation that estimates the time weighted average shear wave velocity at 30 meters depth ($\bar{V}_{s,30}$) based on shallower velocity values. This correlation can be seen on the next page;

$$\ln(\bar{V}_{s,30}) = a + b\ln(\bar{V}_{s,d})$$

where a and b are empirical constants and $\bar{V}_{s,d}$ is the time weighted average shear wave velocity at depth, d. Since the values at 10-meters depth were the most constrained/had the most amount of data pairs, this data was used to estimate $\bar{V}_{s,30}$ using Boore's correlation. The empirical semi-variogram and the semi-variogram modeled with a Gaussian model can be seen in Figure 4.14. The additional graphs for this depth can be seen in the Appendix. The range for this estimate is 90 meters.

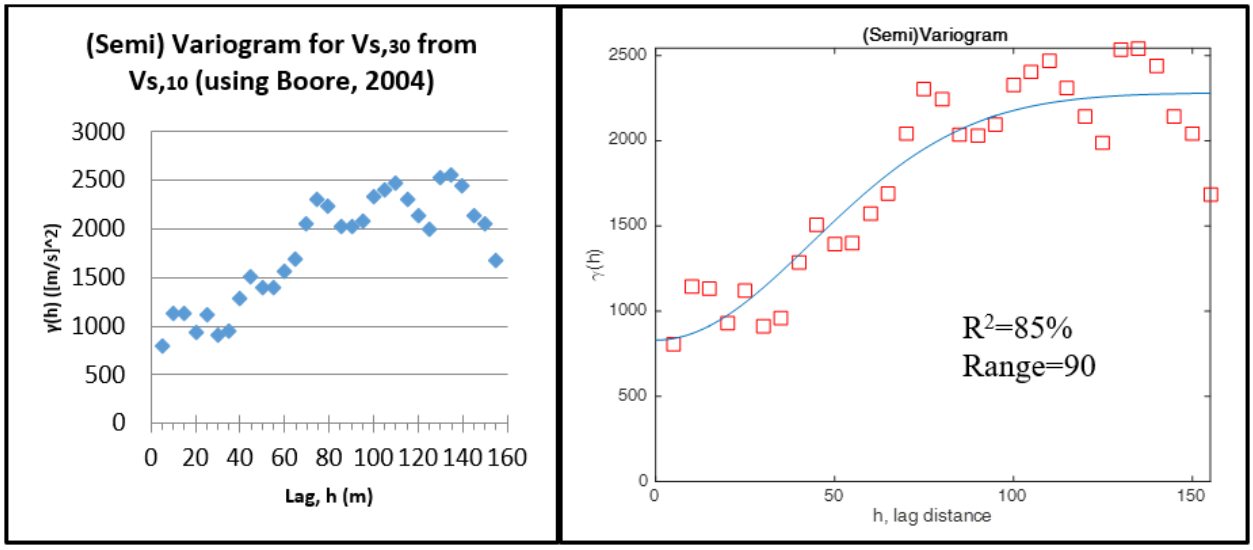


Figure 4.14 Semi-variogram at 30 meters depth from Boore's estimate using velocity values at 10 meters depth; Empirical semi-variogram values (left) and semi-variogram with a Gaussian model (right)

4.5 Applicability and Limitations

4.5.1 Applying Range Values to Other Sites

The range values are designed to fit the Cuesta College site, but the hope is that these values can be applied to other, similar sites as well. This relates primarily to the similarities in the deposition of other sites in comparison to this site. Since the site consists of alluvial deposits overlaying the Franciscan complex, the applicability to other sites with this deposition is the

most appropriate. For other alluvial sites though, these range estimates could be used as estimates or guidelines.

Franciscan mélange which surrounds the site is known to be highly variable. The geologic map (California Conservation and California Geologic Survey, 2010) describes this mélange as a "chaotic mixture of fragmented rock masses". If the alluvium was eroded from this material and the presence of this material at the site indicates higher variability which leads to shorter (more conservative) range values. So even in sites with alluvium overlying other types of rock, these range values could be used as a conservative estimate. Isaaks and Srivastava (1989) also note that it is reasonable to infer the shape of a horizontal semi-variogram from similar data sets. So another option is to sample a few points and to infer the shape of the semi-variogram from this study's more densely sampled area.

It should also be noted that the range values in this study are based exclusively off of the velocity values. And even though the soil type and the velocity values are strongly associated, the similar sites would not necessarily need to have alluvial soils – just velocity values that are similarly distributed. This would involve a comparison between the velocity deposition of alluvial sites compared to the velocity deposition of another site/type of material. This comparison is outside the scope of this study, but the values could be used as guidelines in study's looking at these comparisons.

4.5.2 Reliability of the Range Values

The reliability can be checked by looking at the bias in the survey. The largest bias comes from the array direction. Since this material was deposited by rivers, the material is expected to have anisotropic velocity values (i.e. the cross-stream values are likely to be different than the down-stream values). This anisotropy can lead to changes is the range value (Isaaks and

Srivastava, 1989). Directions that are more similar would have higher range values and vice-versa. Looking at the distribution of alluvium in Figure 1.2 and seeing that Chorro creek runs in an east-west direction just below Cuesta's campus, it appears that this study's survey sampled velocity values in the cross-stream direction. Values sampled in the down-stream direction could give different range values. This anisotropy is not checked in this study and further studies would need to verify the accuracy of the range values that are presented.

Chapter 5: Conclusion

5.1 Summary

The time weighted average Rayleigh wave velocities and shear wave velocities were collected along a 310 meter long profile in alluvial soil. Semi-variograms, covariance functions, and correlograms were created from this data to quantify the spatial continuity. Negative covariance values indicated that a trend could be present in the data, so the data was manipulated in four different ways to try to relieve the data of this apparent trend. However, after modeling all the options with spherical models, exponential models, and Gaussian models, the manipulated data did not appear to model the data any better than the original data. Although stationarity was violated with this data (making the covariance functions and correlograms unusable), intrinsic stationarity was assumed so that the semi-variograms would still give useful information. The spherical and Gaussian models both captured the uncertainty in the data equally well, but the Gaussian model's ranges were more conservative so these were preferred over the spherical models. Ranges of 50 meters for a 5-meter column of soil and approximately 65 meters for a 10- to 15-meter column of soil when Rayleigh waves were calculated. For shear waves, the models showed ranges of 70 meters for a 5-meter column of soil and of 100 meters for a 10- to 15-meter column of soil. Using Boore's (2004) study to extrapolate deeper, the range for a 30-meter column of soil was 90 meters.

The range values for shear waves were consistently higher than those for Rayleigh waves. This implies that the Rayleigh wave values were consistently less similar than compared to the shear wave values. The most likely reason for this is because the Rayleigh wave data was calculated from an approximation involving about 10 layers, whereas the shear wave data was calculated via an inversion process and only 3 to 5 layers. The Rayleigh wave data was strictly

from the dispersion curve, but during the inversion process iterations continue and the amount of layers can be manipulated until the values seem reasonable. This helps to constrain the shear wave velocities and exclude seemingly erroneous values which leads to shear wave values that are more similar and higher range values.

Even though the shear wave values are manipulated during the inversion process, this manipulation considers prior knowledge of the site. The geologic map, nearby borings, and nearby cross section were used to help constrain the inversion process. And, although this creates bias in the values, the goal is to have the bias based on values that are more likely to be present. These expected values which are anticipated from prior information (boring logs, geologic maps, etc.) are assumed to be more accurate than estimates from the dispersion curve. This is partly because the expected shear wave values help to constrain the data and the noise in the dispersion curve can lead to erroneous values that could influence the Rayleigh wave velocities. Additionally, shear wave velocities are used more in practice and the range from these velocities are more applicable than the range from the less used Rayleigh wave velocities.

This gives a range of 70 meters for 5 meter depths, 100 meters for 10 to 15 meters depth, and (using Boore's 2004 study) 90 meters for 30 meters depth. These values are similar to the ranges proposed by the European code (20 to 200 meters) and on the lower, more conservative end of the ranges proposed by the Russian code (100 to 500 meters if powerline installation is ignored). Note that the ranges in these codes are based on the type of structure being built whereas this study's ranges are based on the deposition of soil only.

5.2 Future Research

Recommendations for future research can be grouped into two categories: confirming this study's results and seeing how well these results translate to other soil types/depositions.

Verifying the results can be done at another site or even at the same site using a perpendicular/two-dimensional survey. A different site would allow for a comparison of the range values and a possibility to find useable covariance functions/correlograms; a study at the same site could check for anisotropy (and any corresponding changes in the range due to this anisotropy). A two-dimensional survey requires more intensive analysis (Isaaks and Srivastava, 1989, give a good summary of this) but would allow for multiple directions to be checked and for more data pairs to be analyzed with less survey points. If this approach is taken, it should be noted that Oliver and Webster (2015) advise that only 5 data points are needed in a semi-variogram before the sill. This study used a small sample interval (5 meters) to have better chances of identifying the range and to have more data points, but since the range value is now assumed to be around 90 meters for shear waves, then this only requires a sample interval of 15 meters in order to give 5 points before the sill.

The second option is to model a site with a different soil deposition to see how well the range values match. Since the range values are based on the differences in velocity values, a future study could model a site with similar velocity values to see whether the soil deposition or the velocity values influence the range more. Of course, other sites with vastly different velocity values and deposition could also be modeled, but it seems that these sites would be less likely to show similar range values.

REFERENCES

Achenbach, J. D. (1973). *Wave Propagation in Elastic Solids*. Amsterdam, The Netherlands: North-Holland Publishing.

Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors. *BRRI*, *35*, p. 415–456

Aki, K. (1965). A note on the use of microseisms in determining the shallow structure of the Earth's crust. *Geophysics, 30*, 665-666.

Aki, K., & Richards, P. (2002). *Quantitative Seismology* (2nd ed.). Sausalito, California: University Science Books.

Andrus, R., & Stokoe II, K. (1996). Liquefaction Resistance of Soils from Shear wave Velocity. *J. Geotech. Geoenviron. Eng. Journal of Geotechnical and Geoenvironmental Engineering*, 1015-1025.

Asten, M., & Henstridge, J. (1984). Array estimators and the use of microseisms for reconnaissance of sedimentary basins. *Geophysics, 49*(11), 1828-1837.

Bohling, G. (2005). *Introduction to Geostatistics and Variogram Analysis* [PDF document]. Retrieved from http://people.ku.edu/~gbohling/cpe940/Variograms.pdf

Boore, D. M. (2004). Estimating $\bar{V}_s(30)$ (or NEHRP Site Classes) from Shallow Velocity Models (Depths < 30 m). *Bulletin of the Seismological Society of America, 94*(2), 591-597.

Bracewell, R. N. (1978). *The Fourier Transform and Its Applications.* 2nd ed. New York: McGraw-Hill.

California Conservation and California Geologic Survey [Mark O. Wiegers]. (2010). Geologic Map of the San Luis Obispo 7.5' Quadrangle San Luis Obispo County, California: A Digital Database [map]. 1:24,000. Retrieved from ftp://ftp.consrv.ca.gov/pub/dmg/rgmp/Prelim_geo_pdf/SanLuisObispo24k_preliminary.pdf

Chipping, D. (1987). *The Geology of San Luis Obispo County: A Brief Description and Field Guide*. San Luis Obispo, California: D. H. Chipping.

Clark, I. (1979). *Practical Geostatistics*. London: Applied Science.

EN 1997-2: Eurocode 7: Geotechnical Design – Part 2: Ground Investigation and Testing (English ed.). (2007). European Committee for Standardization.

Engineering Surveys for Construction. Basic Provisions (SNiP 11-02-96 ed.). (1996). Russian Federation.

Facciorusso, J., Uzielli, M., & Mayne, P. (2010). *Spatial characterization of Vs at Amherst NGES site from SCPT using Bayesian kriging*. Conference paper at 2nd International Symposium on Cone Penetration Testing, Huntington Beach, CA.

Foti, S., Lai, C., Rix, G., & Strobbia, C. (2014). *Surface Wave Methods for Near-Surface Site Characterization*. CRC Press.

Geogiga Surface Plus (Version 7.1) [Software]. (2012). Calgary, Alberta, Canada: Geogiga Technology Corporation.

Google Maps. (2015). [Cuesta College, San Luis Obispo, California] [Map]. Retrieved from https://www.google.com/maps/place/Cuesta+College/@35.2249348,-120.1304288,9z/data=!4m2!3m1!1s0x80ece51406b45c83:0xf478e2d29c4fed3a

Hayashi, K. (2015). *Array Microtremor (SPAC); past, present, future & comments regarding InterPacific project*. Unpublished presentation, Post-SSA Forum, California Institute of Technology, Pasadena, CA.

Horike, M. (1985). Inversion of phase velocity of long period microtremors to the S-wave velocity structure down to the basement of urbanized areas. *J Phys Earth, 33*, 59-96.

Hudson, J. A. (1980). *The Excitation and Propagation of Elastic Waves.* Cambridge: Cambridge University Press.

Isaaks, E., & Srivastava, R. (1989). *An Introduction to Applied Geostatistics*. New York: Oxford University Press.

Iqbal, J., Thomasson, J., Jenkins, J., Owens, P., & Whisler, F. (2005). Spatial Variability Analysis of Soil Physical Properties of Alluvial Soils. *Soil Science Society of America Journal, 69*, 1338-1350. doi:10.2136/sssaj2004.0154

Joyner, W. B., R. E. Warrick, & Fumal, T. E. (1981). The effect of Quaternary alluvium on strong ground motion in the Coyote Lake, California, earthquake of 1979, *Bulletin of the Seismolocial Society of America*, *71*, 1333-1349.

Lai, C. G. (2005). Surface waves in dissipative media: Forward and inverse modeling. *Surface Waves in Geomechanics: Direct and Inverse Modelling for Soils and Rocks* (C.G. Lai and K. Wilmanski, eds.). Wien, New York: Springer-Verlag.

Lai, C. G., Foti, S., & Rix, G. J. (2005). Propagation of data uncertainty in surface wave inversion. *J Environ Eng Geophys, 10*(2), 219-228.

Lamb, H. (1904). On the propagation of tremors over the surface of an elastic solid. *Philos Trans Roy Soc Lond Ser A, 203,* 1-42.

Louie, J. N. (2001). Faster, Better: Shear wave Velocity to 100 Meters Depth from Refraction Microtremor Arrays. *Bulletin of the Seismological Society of America,* 347-364.

Malagnini, L., Rovelli, A., Hough, S., & Seeber, L. (1993). Site Amplification Estimates in the Garigliano Valley, Central Italy, Based on Dense Array Measurements of Ambient Noise. *Bulletin of the Seismological Society of America, 83*(6), 1744-1755.

Marosi, K. T., & Hiltunen, D. R. (2004). Characterization of SASW phase angle and phase velocity measurement uncertainty. *Geotech Test J, 27*(2), 205-213.

Matheron, G. (1963). *Les variables régionalisées et leur estimation*. Paris: Masson.

Matheron, G. (1965). Principles of Geostatistics. *Economic Geology, 58,* 1246–1266.

Moss, R. (2008). Quantifying Measurement Uncertainty of Thirty-Meter Shear wave Velocity. *Bulletin of the Seismological Society of America,* 1399-1411.

Oliver, M., & Webster, R. (2015). *Basic Steps in Geostatistics:The Variogram and Kriging*. Springer.

O'Neill, A. (2003). *Full-waveform reflectivity for modelling, inversion and appraisal of seismic surface wave dispersion in shallow site investigations*. University of Western Austrailia.

Parker, R. L. (1977). Understanding inverse theory. *Ann Rev Earth Planet Sci, 5,* 35-64.

Rayleigh, J. (1877). *The Theory of Sound* (1st ed.). London: Macmillan and Co.

Rix, G., & Leipski, E. (1991). Accuracy and Resolution of Surface Wave Inversion. *Recent advances in instrumentation, data acquisition and testing in soil dynamics: Am. Soc. Civil Eng.*, 17–32.

Santamarina, J., Klein, K., & Fam, M. (2001). *Soils and Waves*. Chichester: J. Wiley & Sons.

Schwanghart, W. (2010) *Variogramfit* (Version 1.5) [MATLAB code]. Available at http://www.mathworks.com/matlabcentral/fileexchange/25948-variogramfit (Accessed July 2015)

Shekhar, S. (2008). Kriging. In H. Xiong (Ed.), *Encyclopedia of GIS*. New York: Springer.

Stein, S., & Stein, J. (2014). Nature Bats Last. In *Playing Against Nature: Integrating Science and Economics to Mitigate Natural Hazards in an Uncertain World*. John Wiley & Sons.

Stover, Christopher & Weisstein, Eric W. "Fréchet Derivative." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/FrechetDerivative.html

Strang, G. (1980). *Linear Algebra and Its Applications* (2nd ed.). New York: Academic Press.

Telford, W., & Geldart, L. (1990). *Applied Geophysics* (2nd ed.). Cambridge: Cambridge University Press.

Tezcan, S., Keceli, A., & Ozdemir, Z. (2006). Allowable Bearing Capacity of Shallow Foundations Based on Shear Wave Velocity. *Geotechnical and Geological Engineering Geotech Geol Eng, 24*, 203-218.

Thitimakorn, T. (2013). Development of a NEHRP site classification map of Chiang Mai city, Thailand, based on shear wave velocity using the MASW technique. *J. Geophys. Eng. Journal of Geophysics and Engineering*, 045007-045007.

Tuomi, K. E., & Hiltunen, D. R. (1996). Reliability of the SASW method for determination of the shear modulus of soils. *Proceedings of Uncertainty in Geologic Environment: From Theory to Practice.* American Society of Civil Engineers. Shacklford, Reston, Virginia. 1125-1237.

Tokimatsu, K. (1995). Geotechnical Site Characterization using Surface Waves. *Proceedings of 1st International Conference on Earthquake Geotechnical Engineering, 3*, 1333-1368.

Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics, 74*(6), WCC127-WCC152. doi:10.1190/1.3238367

Webster, R. (2000). Is soil variation random? *Geoderma, 97*, 149–163.

Webster, R., & Oliver, M. A. (1992). Sample adequately to estimate variograms of soil properties. *Journal of Soil Science, 40,* 493–496.

Wills, C. J., & Clahan, C. B. (2006). Developing a map of geologically defined site-condition categories for California, *Bull. Seismo. Soc. Am.*, *96*(4A):1483–1501.

Appendix A: Earth Systems Pacific's Boring Logs for the Cuesta College Sewer Line and Pipe Bridge Replacement Project (2010)

Earth Systems Pacific

BORING LOG LEGEND

SAMPLE / SUBSURFACE WATER SYMBOLS	GRAPH. SYMBOL
CALIFORNIA MODIFIED	
STANDARD PENETRATION TEST (SPT)	
SHELBY TUBE	
BULK	
SUBSURFACE WATER DURING DRILLING	
SUBSURFACE WATER AFTER DRILLING	

SOIL CLASSIFICATION SYSTEM

MAJOR DIVISIONS	GROUP SYMBOL	TYPICAL DESCRIPTIONS	GRAPH. SYMBOL
COARSE GRAINED SOILS (MORE THAN HALF OF MATERIAL IS TESTED OR JUDGED TO BE LARGER THAN #200 SIEVE SIZE)	GW	WELL GRADED GRAVELS, GRAVEL-SAND MIXTURES, LITTLE OR NO FINES	
	GP	POORLY GRADED GRAVELS, GRAVEL-SAND MIXTURES, LITTLE OR NO FINES	
	GM	SILTY GRAVELS, GRAVEL-SAND-SILT MIXTURES, NON-PLASTIC FINES	
	GC	CLAYEY GRAVELS, GRAVEL-SAND-CLAY MIXTURES, PLASTIC FINES	
	SW	WELL GRADED SANDS, GRAVELLY SANDS, LITTLE OR NO FINES	
	SP	POORLY GRADED SANDS, GRAVELLY SANDS, LITTLE OR NO FINES	
	SM	SILTY SANDS, SAND-SILT MIXTURES, NON-PLASTIC FINES	
	SC	CLAYEY SANDS, SAND-CLAY MIXTURES, PLASTIC FINES	
FINE GRAINED SOILS (HALF OR MORE OF MATERIAL IS TESTED OR JUDGED TO BE SMALLER THAN #200 SIEVE SIZE)	ML	INORGANIC SILTS AND VERY FINE SANDS, SILTY, CLAYEY FINE SANDS, CLAYEY SILTS WITH SLIGHT PLASTICITY	
	CL	INORGANIC CLAYS OF LOW TO MEDIUM PLASTICITY, GRAVELLY CLAYS, SANDY CLAYS, SILTY CLAYS, LEAN CLAYS	
	OL	ORGANIC SILTS AND ORGANIC SILTY CLAYS OF LOW PLASTICITY	
	MH	INORGANIC SILTS, MICACEOUS OR DIATOMACEOUS FINE SANDY, SILTY SOILS, ELASTIC SILTS	
	CH	INORGANIC CLAYS OF HIGH PLASTICITY, FAT CLAYS	
	OH	ORGANIC CLAYS OF MEDIUM TO HIGH PLASTICITY, ORGANIC SILTS	
	PT	PEAT AND OTHER HIGHLY ORGANIC SOILS	

OBSERVED MOISTURE CONDITION

DRY	SLIGHTLY MOIST	MOIST	VERY MOIST	WET
LITTLE/NO MOISTURE	JUDGED BELOW OPTIMUM	JUDGED ABOUT OPTIMUM	JUDGED OVER OPTIMUM	SATURATED

TYPICAL CONSISTENCY

COARSE GRAINED SOILS			FINE GRAINED SOILS		
BLOWS/FOOT		DESCRIPTIVE TERM	BLOWS/FOOT		DESCRIPTIVE TERM
SPT	CA SAMPLER		SPT	CA SAMPLER	
0-10	0-16	LOOSE	0-2	0-3	VERY SOFT
11-30	17-50	MEDIUM DENSE	3-4	4-7	SOFT
31-50	51-83	DENSE	5-8	8-13	MEDIUM STIFF
OVER 50	OVER 83	VERY DENSE	9-15	14-25	STIFF
			16-30	26-50	VERY STIFF
			OVER 30	OVER 50	HARD

GRAIN SIZES

U.S. STANDARD SERIES SIEVE				CLEAR SQUARE SIEVE OPENING			
# 200	# 40	# 10	# 4	3/4"	3"	12"	
SILT & CLAY	SAND			GRAVEL		COBBLES	BOULDERS
	FINE	MEDIUM	COARSE	FINE	COARSE		

TYPICAL ROCK HARDNESS

MAJOR DIVISIONS	TYPICAL DESCRIPTIONS
EXTREMELY HARD	CORE, FRAGMENT, OR EXPOSURE CANNOT BE SCRATCHED WITH KNIFE OR SHARP PICK; CAN ONLY BE CHIPPED WITH REPEATED HEAVY HAMMER BLOWS
VERY HARD	CANNOT BE SCRATCHED WITH KNIFE OR SHARP PICK; CORE OR FRAGMENT BREAKS WITH REPEATED HEAVY HAMMER BLOWS
HARD	CAN BE SCRATCHED WITH KNIFE OR SHARP PICK WITH DIFFICULTY (HEAVY PRESSURE); HEAVY HAMMER BLOW REQUIRED TO BREAK SPECIMEN
MODERATELY HARD	CAN BE GROOVED 1/16 INCH DEEP BY KNIFE OR SHARP PICK WITH MODERATE OR HEAVY PRESSURE; CORE OR FRAGMENT BREAKS WITH LIGHT HAMMER BLOW OR HEAVY MANUAL PRESSURE
SOFT	CAN BE GROOVED OR GOUGED EASILY BY KNIFE OR SHARP PICK WITH LIGHT PRESSURE, CAN BE SCRATCHED WITH FINGERNAIL; BREAKS WITH LIGHT TO MODERATE MANUAL PRESSURE
VERY SOFT	CAN BE READILY INDENTED, GROOVED OR GOUGED WITH FINGERNAIL, OR CARVED WITH KNIFE; BREAKS WITH LIGHT MANUAL PRESSURE

TYPICAL ROCK WEATHERING

MAJOR DIVISIONS	TYPICAL DESCRIPTIONS
FRESH	NO DISCOLORATION, NOT OXIDIZED
SLIGHTLY WEATHERED	DISCOLORATION OR OXIDATION IS LIMITED TO SURFACE OF, OR SHORT DISTANCE FROM; SOME FRACTURES PRESENT; FELDSPAR CRYSTALS ARE DULL
MODERATELY WEATHERED	DISCOLORATION OR OXIDATION EXTENDS FROM FRACTURES, USUALLY THROUGHOUT; Fe-Mg MINERALS ARE "RUSTY"; FELDSPAR CRYSTALS ARE "CLOUDY"
INTENSELY WEATHERED	DISCOLORATION OR OXIDATION THROUGHOUT; FELDSPAR AND Fe-Mg MINERALS ARE ALTERED TO CLAY TO SOME EXTENT OR CHEMICAL ALTERATION PRODUCES IN SITU DISAGGREGATION
DECOMPOSED	DISCOLORATION OR OXIDATION THROUGHOUT, BUT RESISTANT MINERALS SUCH AS QUARTZ MAY BE UNALTERED; FELDSPAR AND Fe-Mg MINERALS ARE COMPLETELY ALTERED TO CLAY

Earth Systems Pacific

LOGGED BY: A. Hinkle
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 1
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 08/07/08

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0			4.5" AC OVER 4.0" AGGREGATE BASE					
1	CH		SANDY FAT CLAY: dark grey brown, hard, moist (alluvium)					
5			coarse gravel	5.0–6.5	Ring	98.8	25.5	5 17 24
10				10.0–11.5	Ring	105.2	20.2	11 21 29
11			light brown, very stiff					
			End of Boring @ 11.5' No subsurface water encountered					

LEGEND: Ring Sample ◯ Grab Sample ▭ Shelby Tube Sample SPT

NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Earth Systems Pacific

LOGGED BY: A. Hinkle
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 2
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 08/07/08

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0			5.0" AC OVER 4.0" AGGREGATE BASE					
1–4	CH		SANDY FAT CLAY: olive brown, very stiff, moist (alluvium)					
4–7	CL		SANDY LEAN CLAY: olive brown, very stiff, moist (residual soil)	5.0–6.5	▩	119.4	13.4	23 29 50
7–11			CLAYSTONE: olive brown, soft, moist (Franciscan formation)	10.0–11.0	●			30 50/2.0"
11			End of Boring @ 11.0' No subsurface water encountered					

LEGEND: ▩ Ring Sample ◯ Grab Sample ▭ Shelby Tube Sample ● SPT

NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Earth Systems Pacific

LOGGED BY: A. Hinkle
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 3
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 08/07/08

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0			4.5" AC OVER 4.0" AGGREGATE BASE					
1	CL		SANDY LEAN CLAY WITH GRAVEL: red brown, medium stiff, moist, coarse gravel (alluvium)	1.0–2.0	○			
3	CL		SANDY LEAN CLAY: dark red brown, hard, moist (residual soil)	3.0–4.0	○			
5				5.0–6.5	▭	107.9	27.9	6 8 50
7			CLAYSTONE: light brown, soft, moist, fractured, coarse gravel (Franciscan formation)	7.0–8.0	○			
10				10.0–10.5	▭	NO RETURN		50/3.0"
11			End of Boring @ 10.5' No subsurface water encountered					

LEGEND: ▭ Ring Sample ○ Grab Sample ▭ Shelby Tube Sample ⊛ SPT

NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Earth Systems Pacific

LOGGED BY: D. Burns
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 4
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 04/09/10

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0			3.5" AC OVER 13.5" AGGREGATE BASE					
2	CH		SANDY FAT CLAY: dark grey brown, stiff, very moist (olluvium)	1.5–4.0	○			
5				5.0–6.5	▩	95.6	15.8	10 11 23
6			very stiff					
10				10.0–11.5	▩	81.1	38.4	9 18 20
12			End of Boring @ 11.5' No subsurface water encountered					

LEGEND: ▩ Ring Sample ○ Grab Sample ▭ Shelby Tube Sample ⊛ SPT

NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Earth Systems Pacific

LOGGED BY: D. Burns
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 5
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 04/09/10

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0			3.0" AC OVER 7.0" AGGREGATE BASE OVER 3.0" AC OVER 4.0" AGGREGATE BASE					
2	CH		SANDY FAT CLAY: dark grey brown, stiff, very moist (alluvium)					
4			coarse gravel					
5			CLAYSTONE: light brown, very soft, moist (Franciscan formation)	5.0–5.5	SPT			50/5.0"
6				4.0–7.0	Grab			
8			soft					
10				10.0–10.0	SPT	NO RETURN		50/1.0"
12			light grey					
15				15.0–15.0	SPT			50/1.0"
			End of Boring @ 15.0' No subsurface water encountered					

LEGEND: Ring Sample ◯ Grab Sample ▭ Shelby Tube Sample SPT

NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Earth Systems Pacific

LOGGED BY: D. Burns
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 6
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 04/09/10

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0	CH		SANDY FAT CLAY: dark grey brown, stiff, moist (alluvium)					
2			very moist					
4	CH		SANDY FAT CLAY: olive/grey mottled, very stiff, very moist (residual soil)	5.0–6.5	Ring	79.6	37.2	4 9 17
10				10.0–11.5	Ring	98.4	23.7	18 24 34
11			End of Boring @ 11.5' No subsurface water encountered					

LEGEND: Ring Sample ◯ Grab Sample ▭ Shelby Tube Sample SPT

NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Earth Systems Pacific

LOGGED BY: D. Burns
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 7
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 04/09/10

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0	SC		CLAYEY SAND WITH GRAVEL: brown, medium dense, moist (fill)					
2	CH		SANDY FAT CLAY: dark grey brown, stiff, very moist (alluvium)					
5				5.0–6.5	Ring	92.1	20.0	8, 11, 9
7	CH		SANDY FAT CLAY: olive brown, very stiff, moist (residual soil)					
10				10.0–11.5	Ring	72.1	32.3	11, 16, 22
12			End of Boring @ 11.5' No subsurface water encountered					

LEGEND: Ring Sample ◯ Grab Sample ▭ Shelby Tube Sample SPT

NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Earth Systems Pacific

LOGGED BY: D. Burns
DRILL RIG: Mobile B-53
AUGER TYPE: 8" Hollow Stem

Boring No. 8
PAGE 1 OF 1
JOB NO.: SL-15726-SB
DATE: 04/09/10

CUESTA COLLEGE SEWER LINE AND PIPE BRIDGE REPLACEMENT
Chorro Valley and Cuesta College Roads
San Luis Obispo, California

DEPTH (feet)	USCS CLASS	SYMBOL	SOIL DESCRIPTION	SAMPLE DATA: INTERVAL (feet)	SAMPLE TYPE	DRY DENSITY (pcf)	MOISTURE (%)	BLOWS PER 6 IN.
0	CH		SANDY FAT CLAY: dark grey brown, stiff, very moist (alluvium)					
5				5.0–6.5	Ring	94.1	27.4	8, 18, 35
6			SANDY LEAN CLAY: light olive, stiff, moist (Residual Soil)					
10				10.0–11.5	SPT			11, 11, 12
11			End of Boring @ 11.5' No subsurface water encountered					

LEGEND: Ring Sample ◯ Grab Sample ▭ Shelby Tube Sample SPT

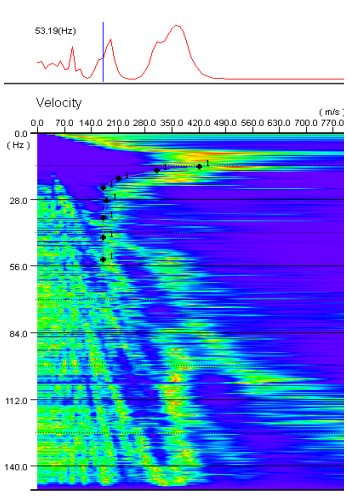
NOTE: This log of subsurface conditions is a simplification of actual conditions encountered. It applies at the location and time of drilling. Subsurface conditions may differ at other locations and times.

Appendix B: SPAC Results for Each Survey

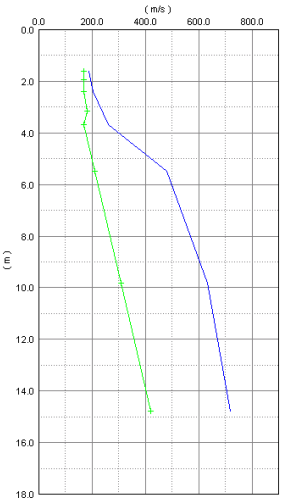
Each survey is referenced based on the center of the array in relation to the reference stake (see Figure 3.2) where northern surveys are negative values and southern are positive values. For example, the array whose center is 55 meters south of the reference stake is denoted as the 55 meter survey. In the following figures, V_R represents the Rayleigh wave velocity and V_S represents the shear wave velocity. The figures below are ordered, from left to right, as the V_R dispersion curve, the V_R profile, the V_S profile, the inversion dispersion curve, and the inversion error iterations.

-20 Meter Survey

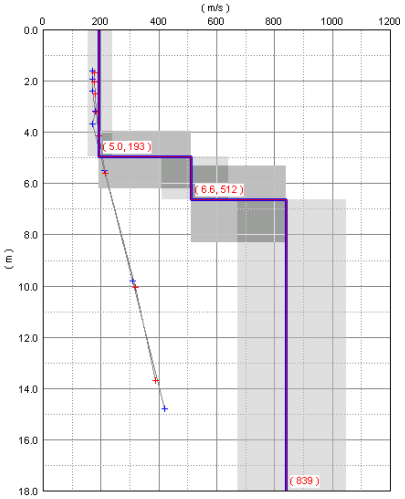
V_R Dispersion Curve



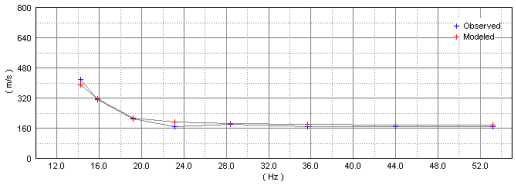
V_R Profile



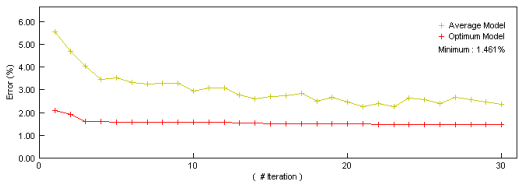
V_S Profile



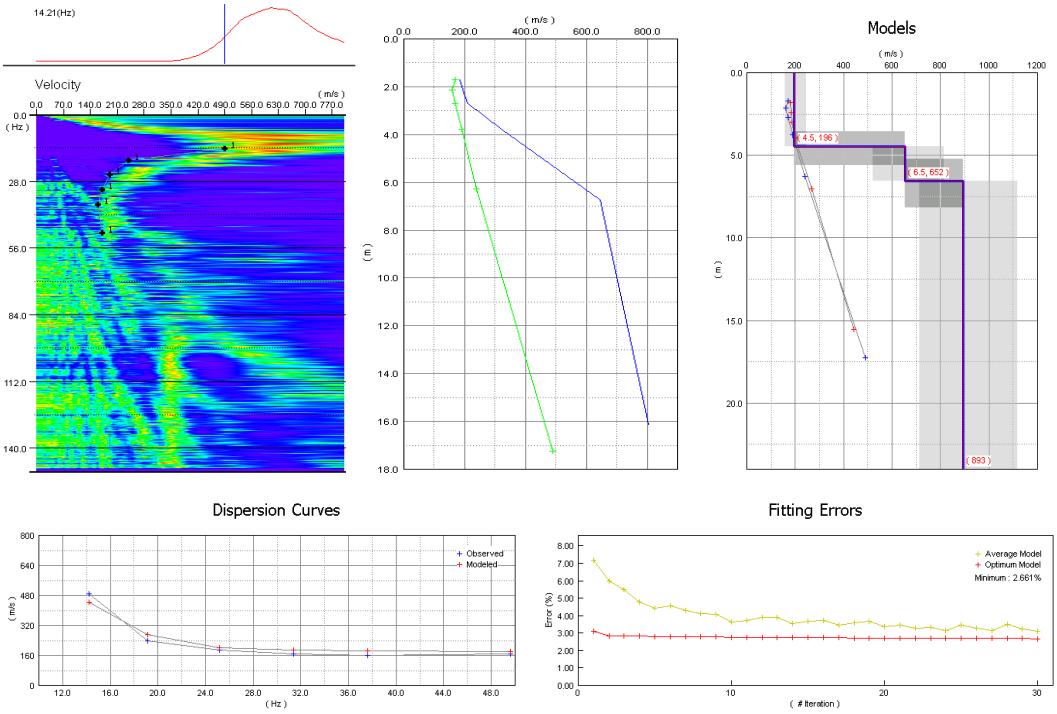
Dispersion Curves



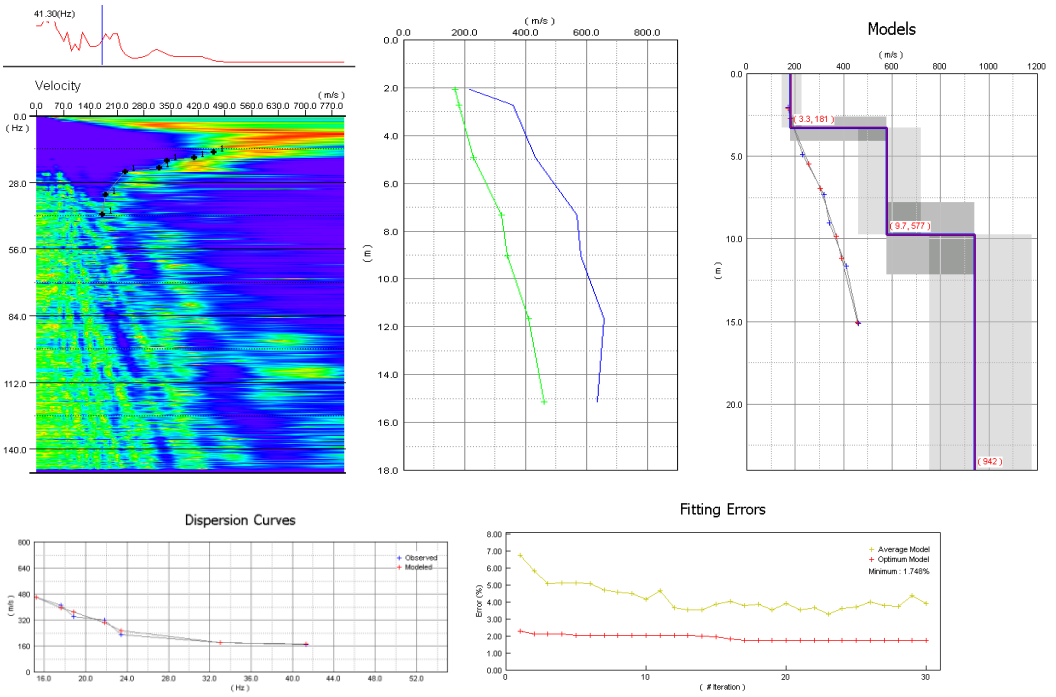
Fitting Errors



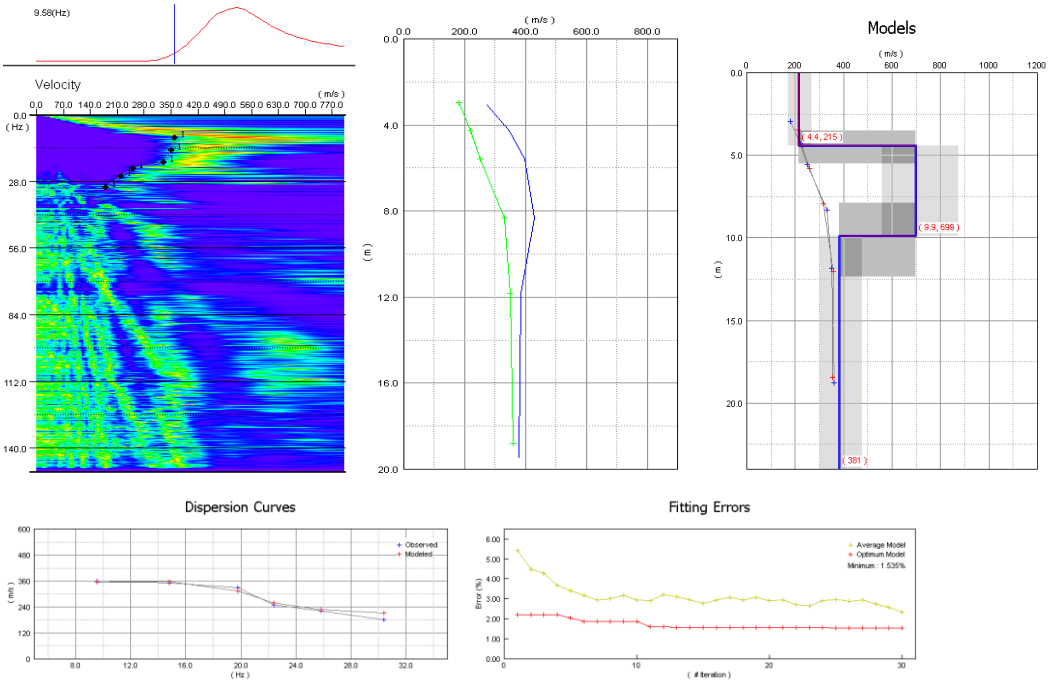
-15 Meter Survey



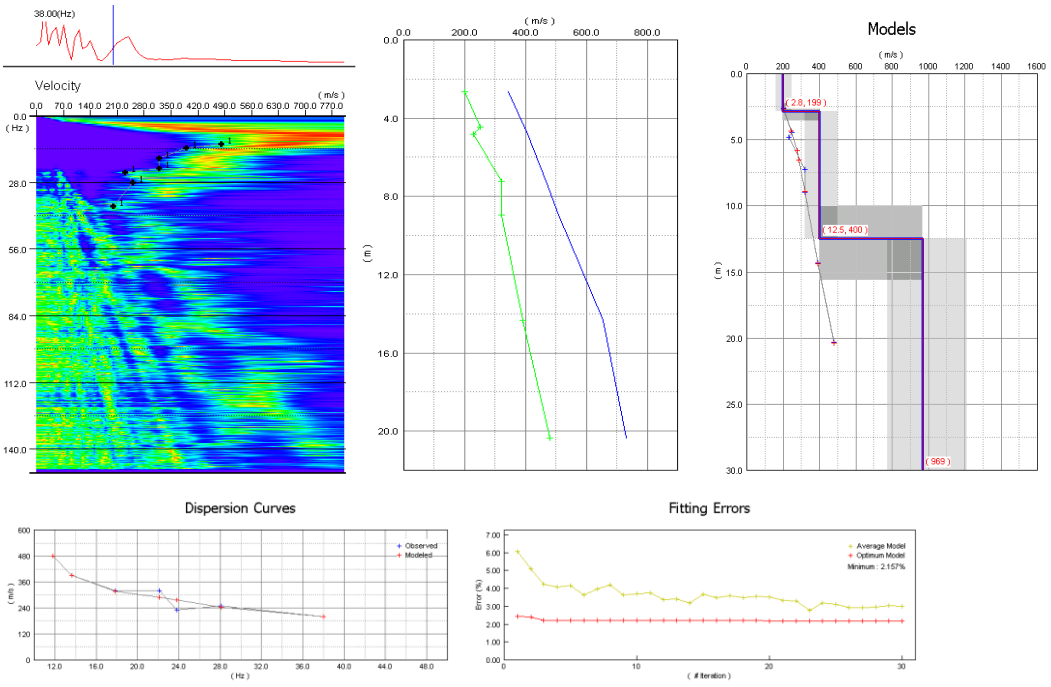
-10 Meter Survey



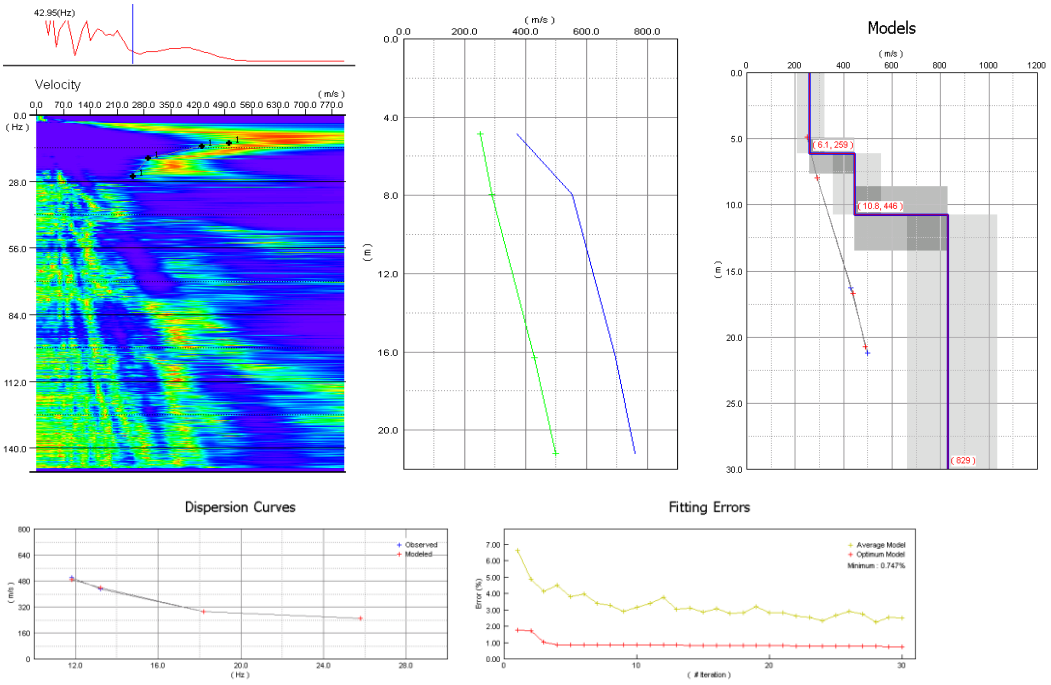
-5 Meter Survey



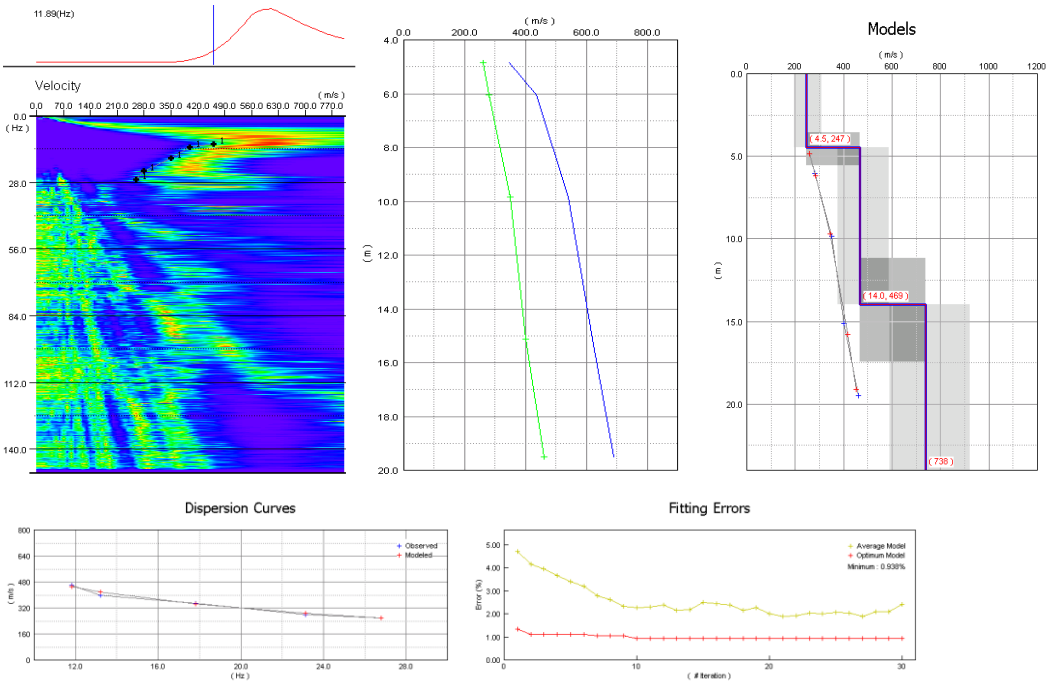
0 Meter Survey



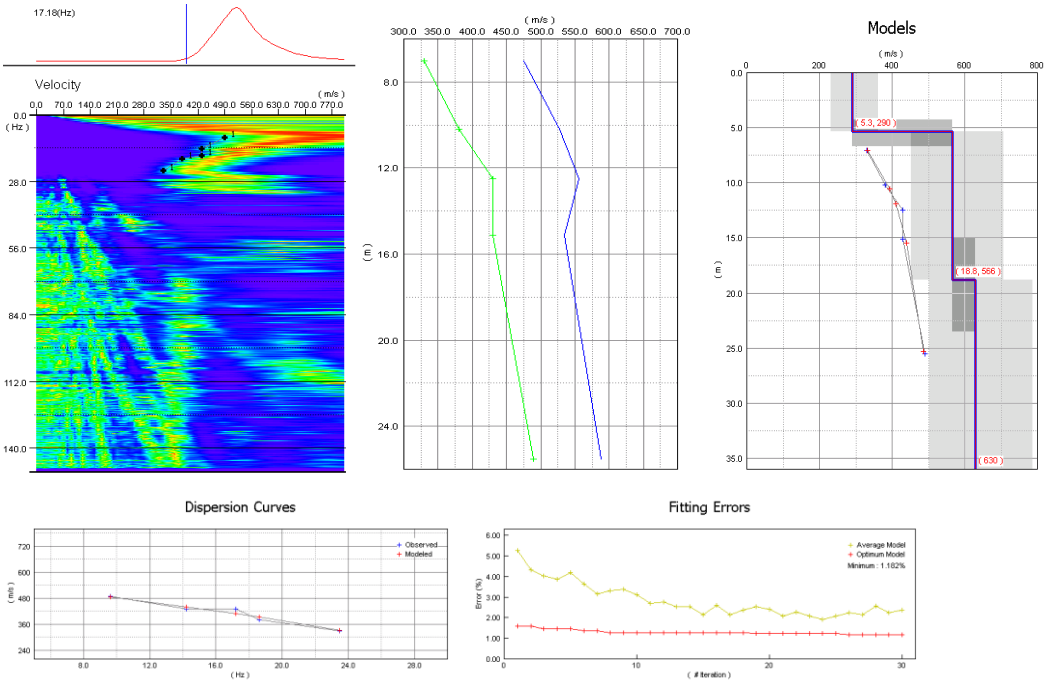
5 Meter Survey



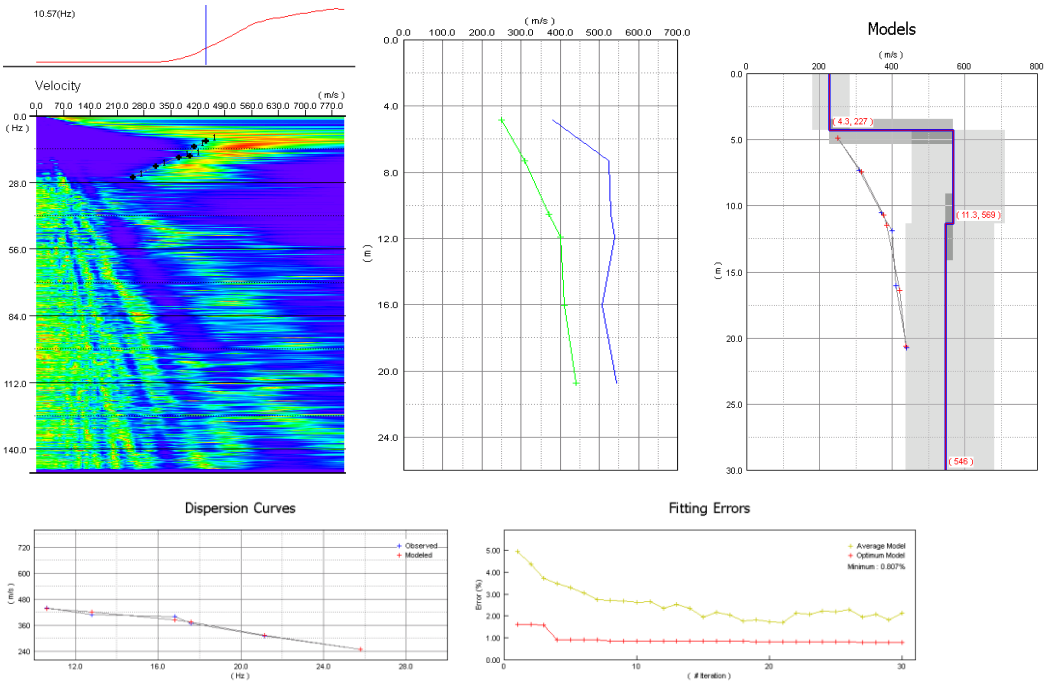
10 Meter Survey



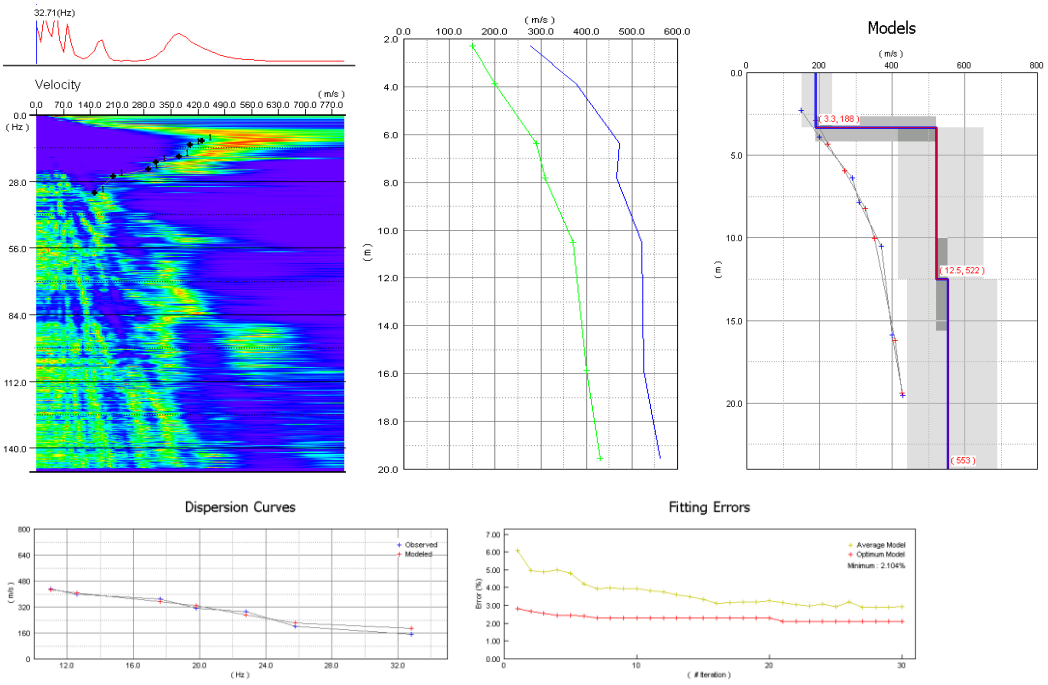
15 Meter Survey



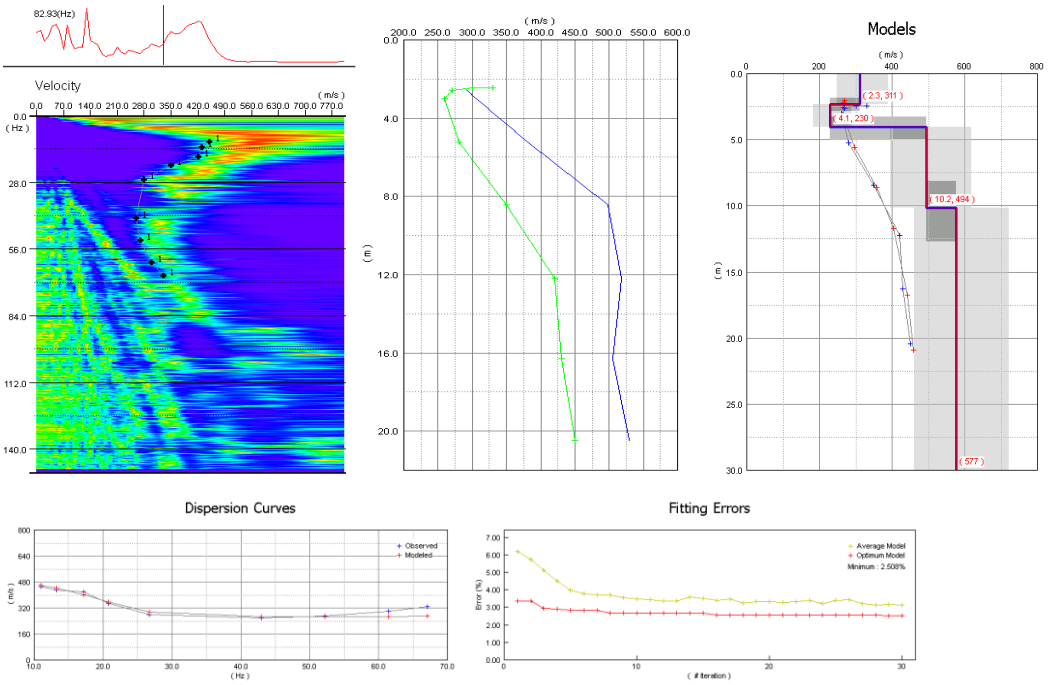
20 Meter Survey



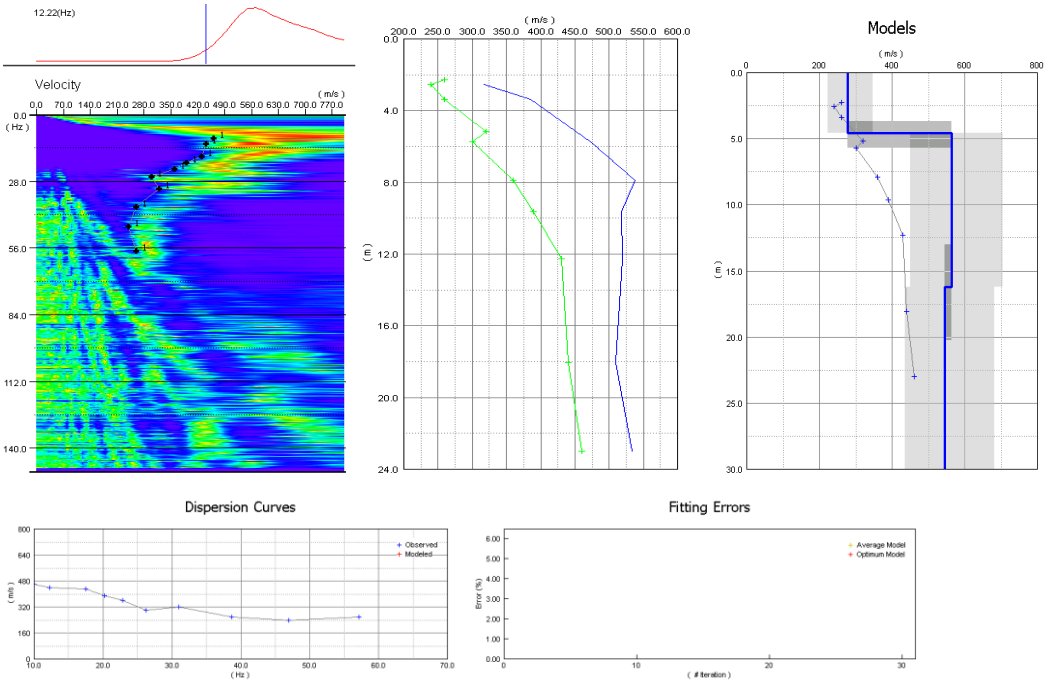
25 Meter Survey



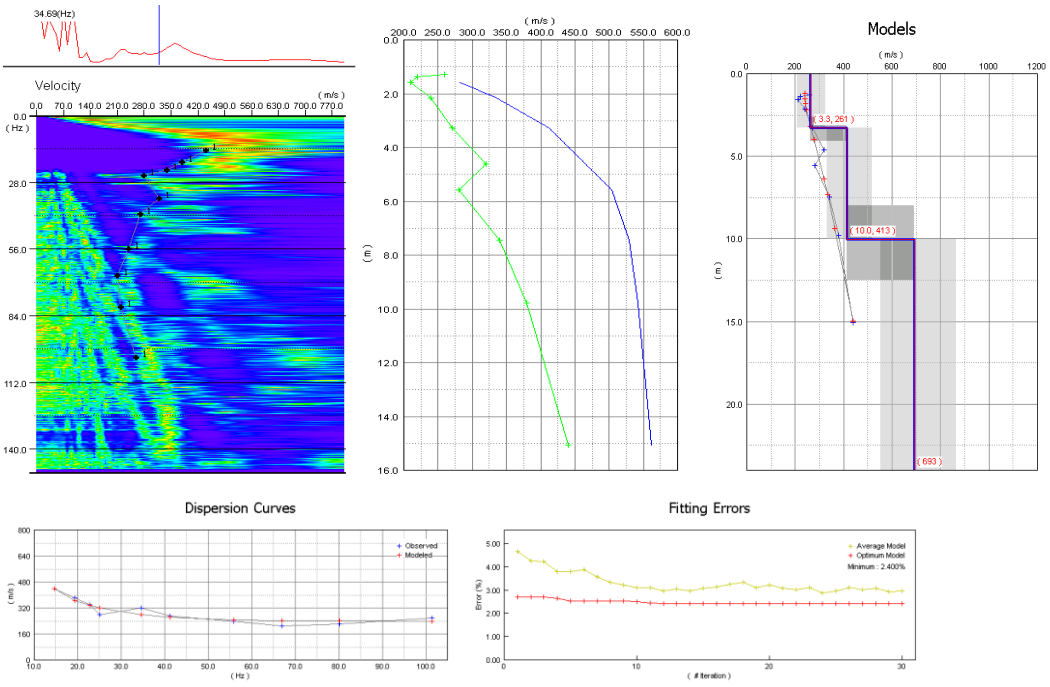
30 Meter Survey



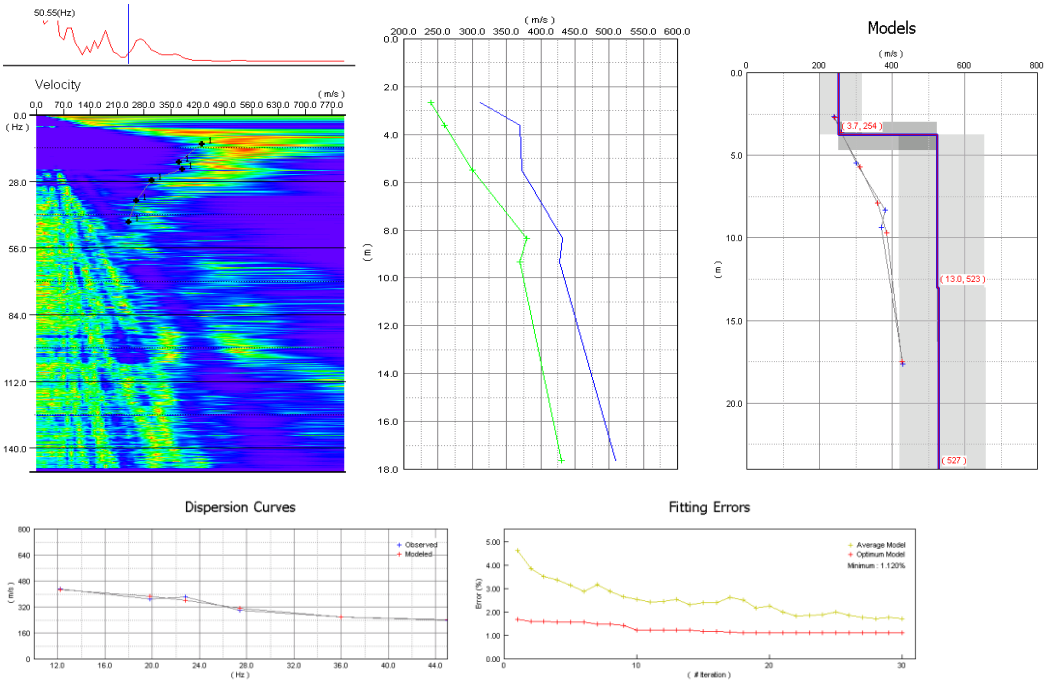
35 Meter Survey



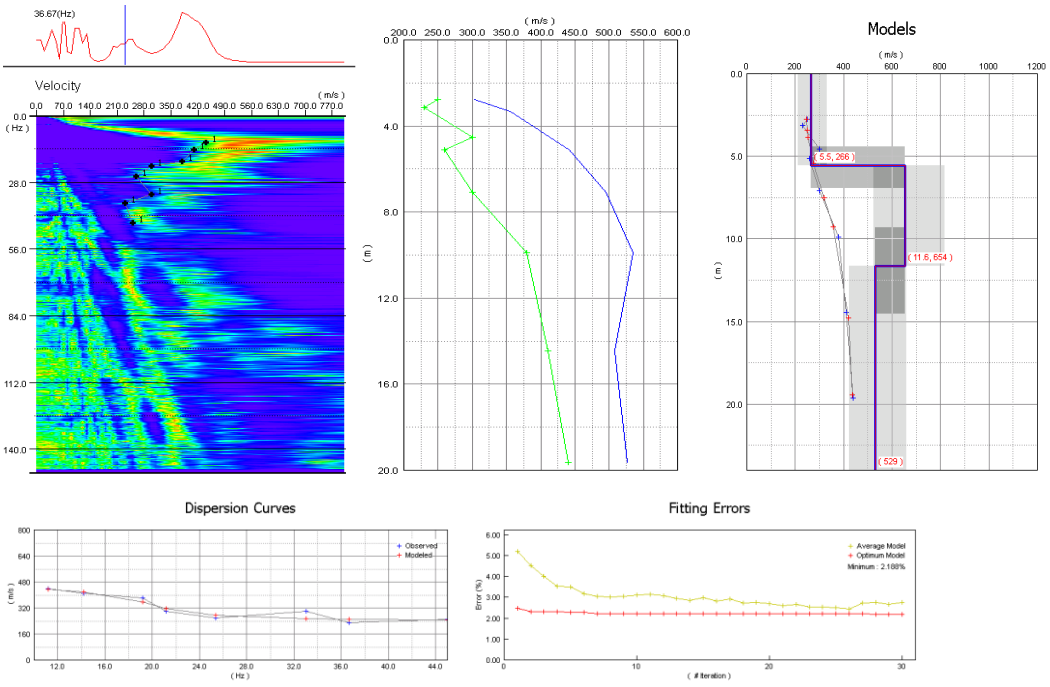
40 Meter Survey



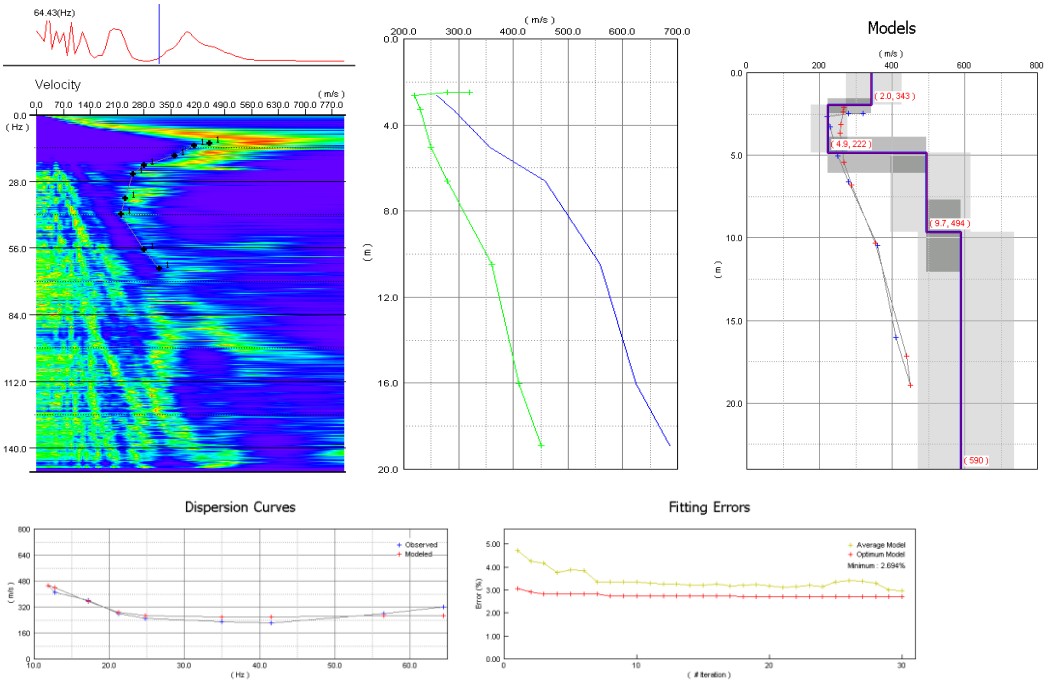
45 Meter Survey



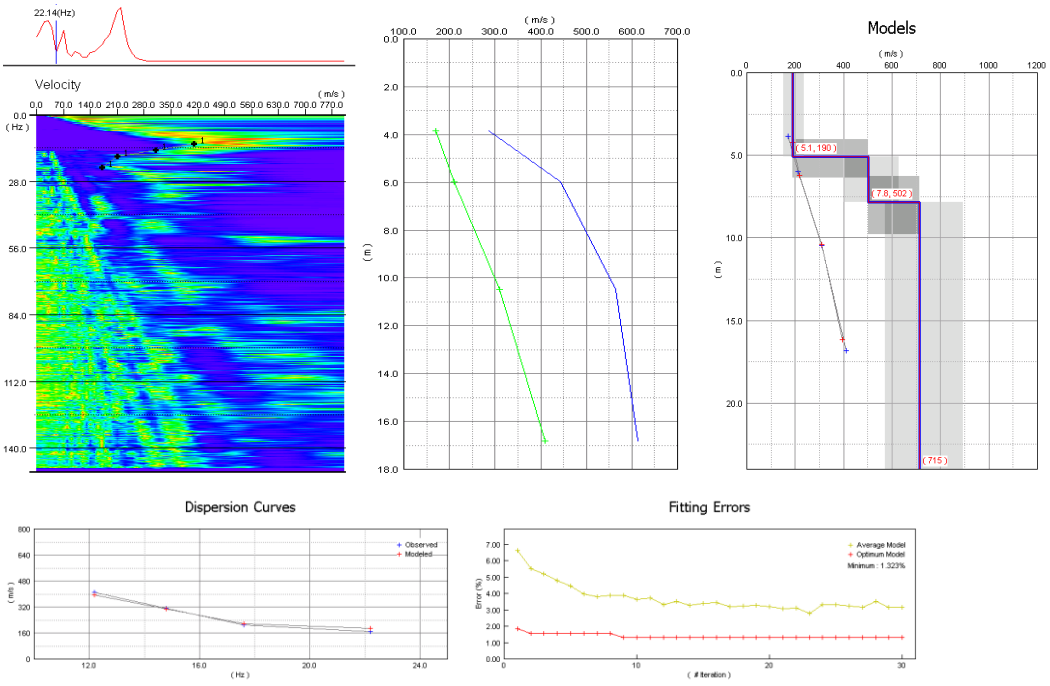
50 Meter Survey



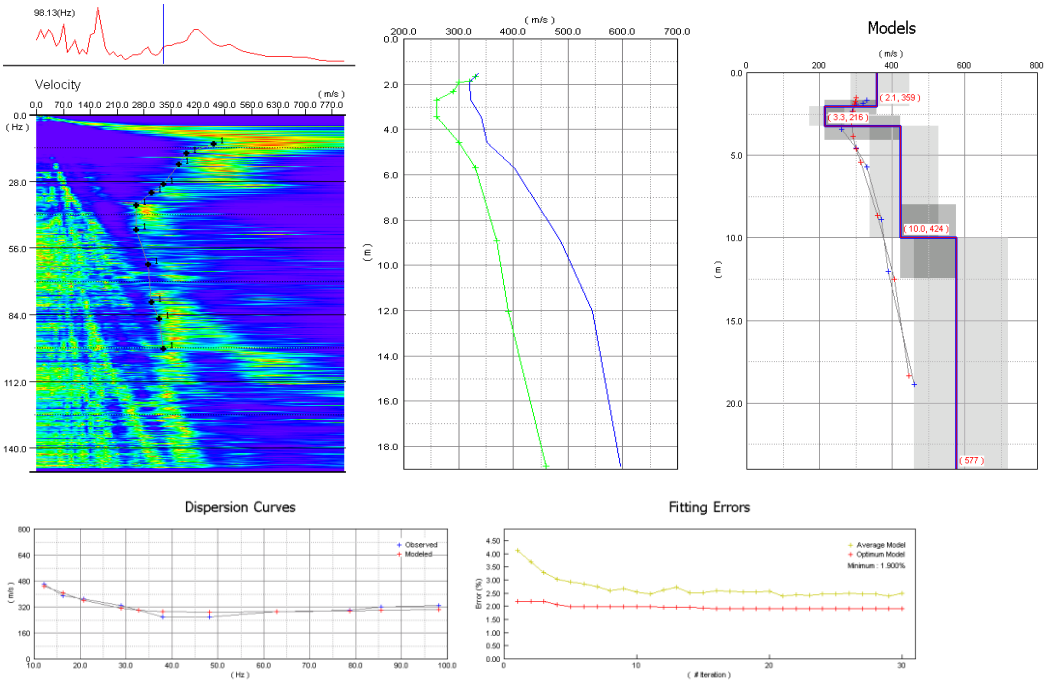
55 Meter Survey



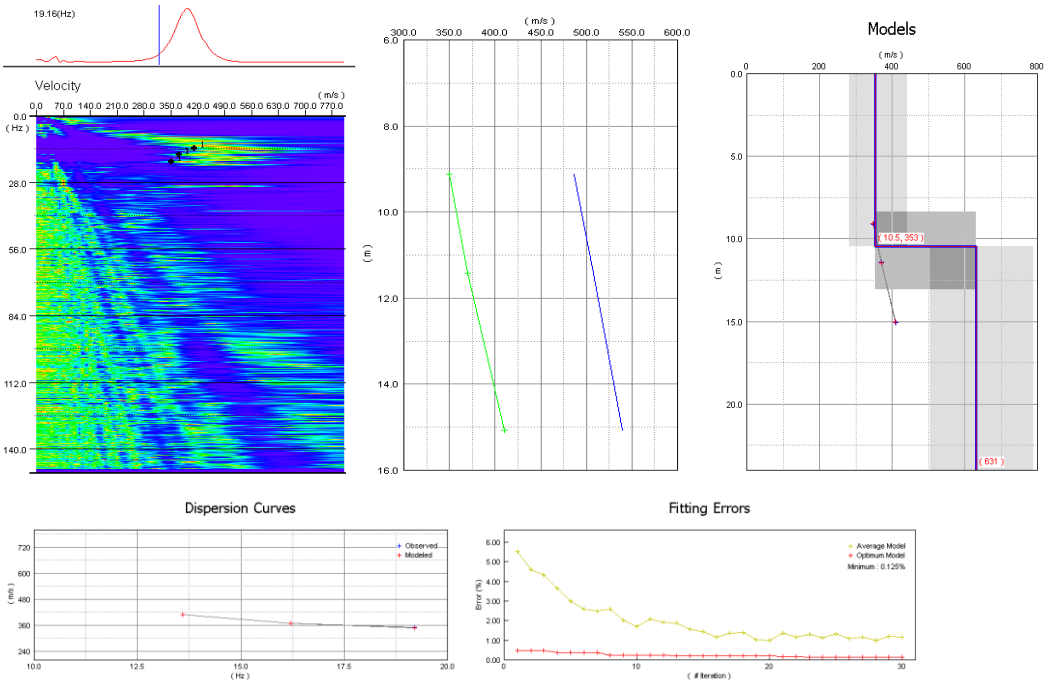
60 Meter Survey



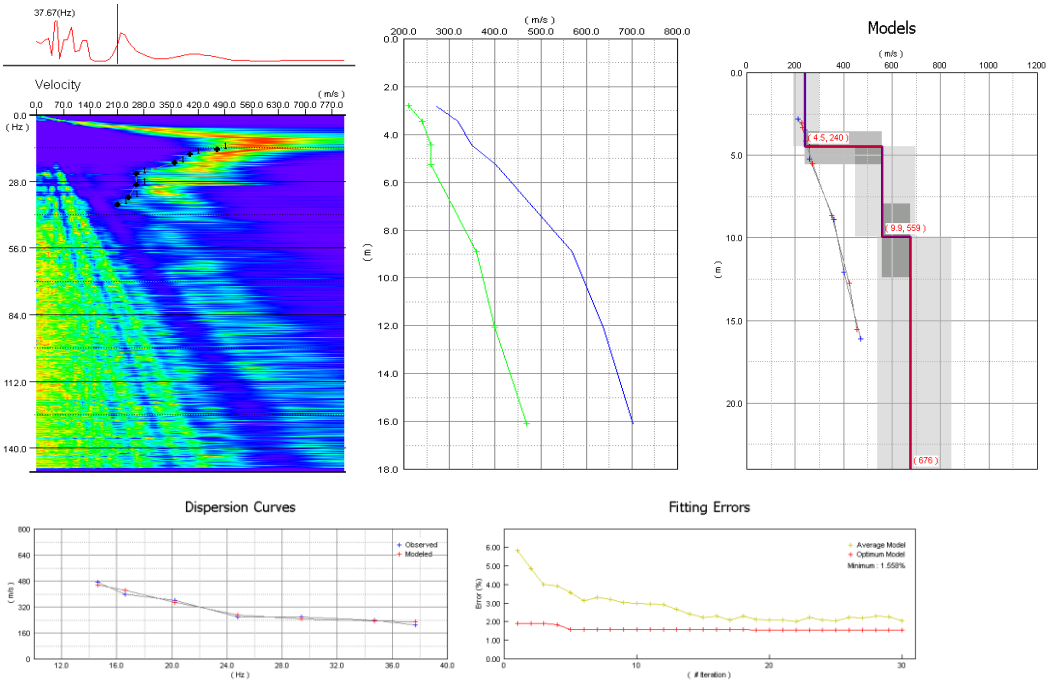
65 Meter Survey



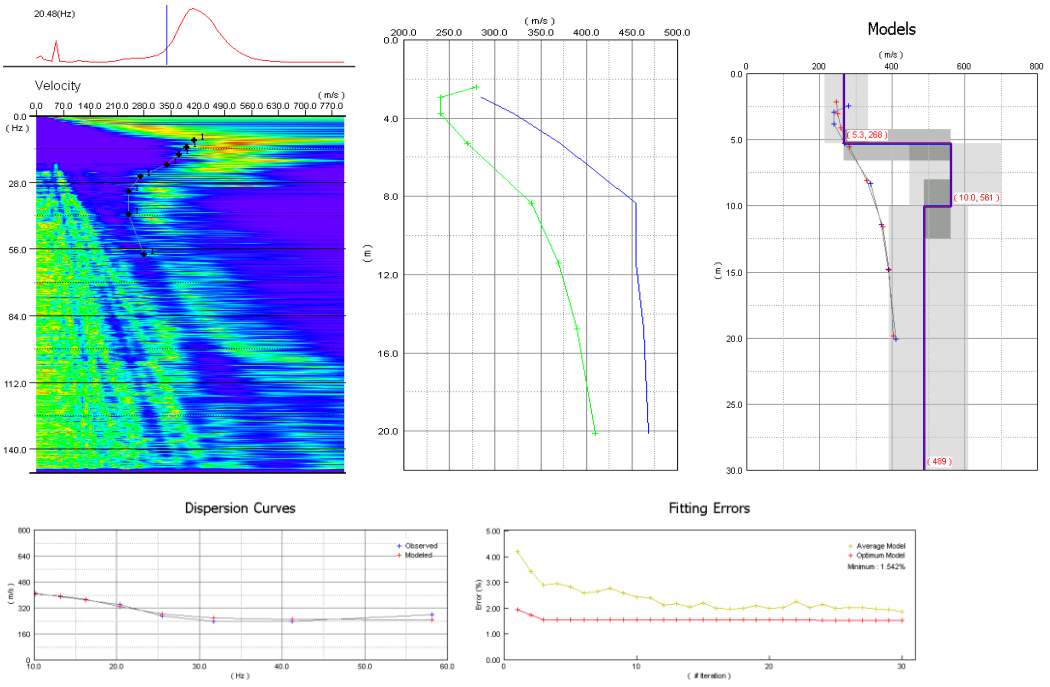
70 Meter Survey



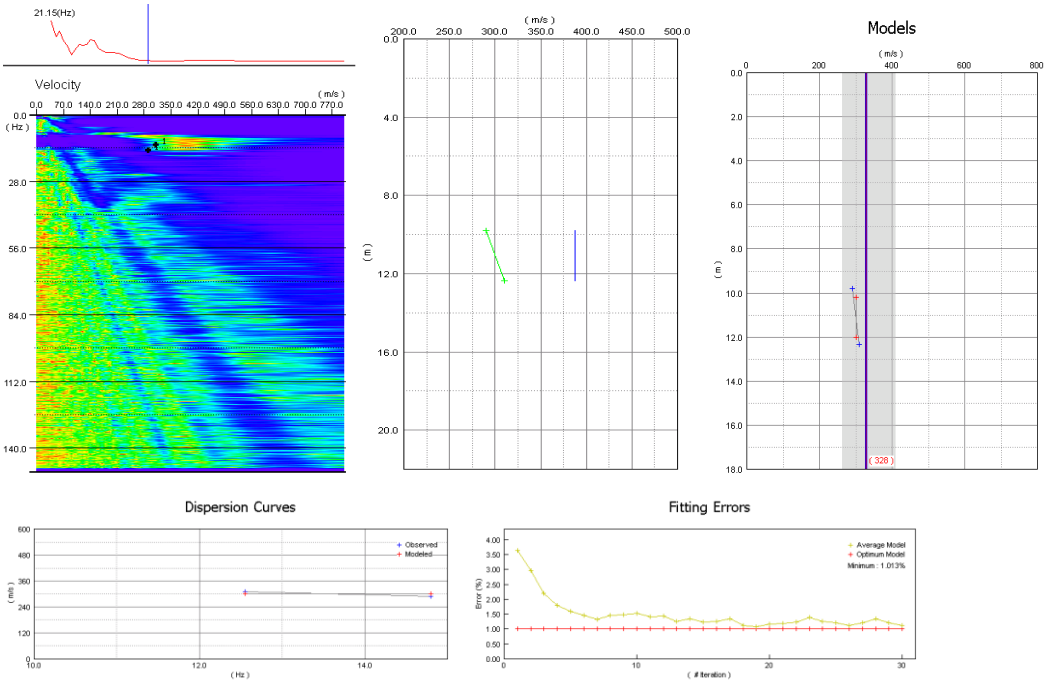
75 Meter Survey



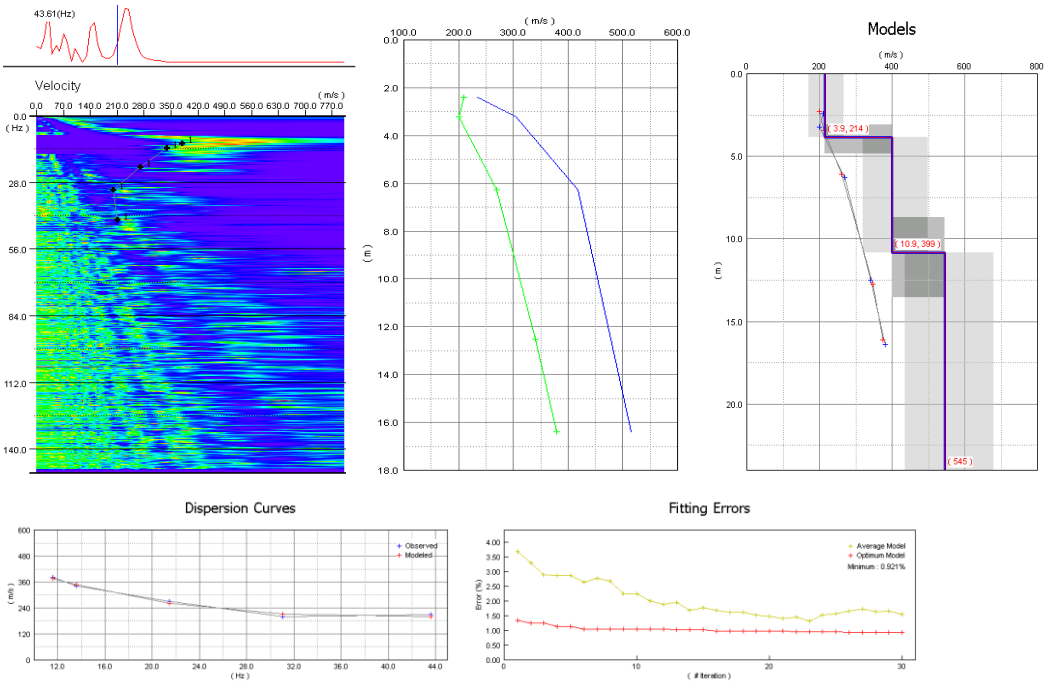
80 Meter Survey



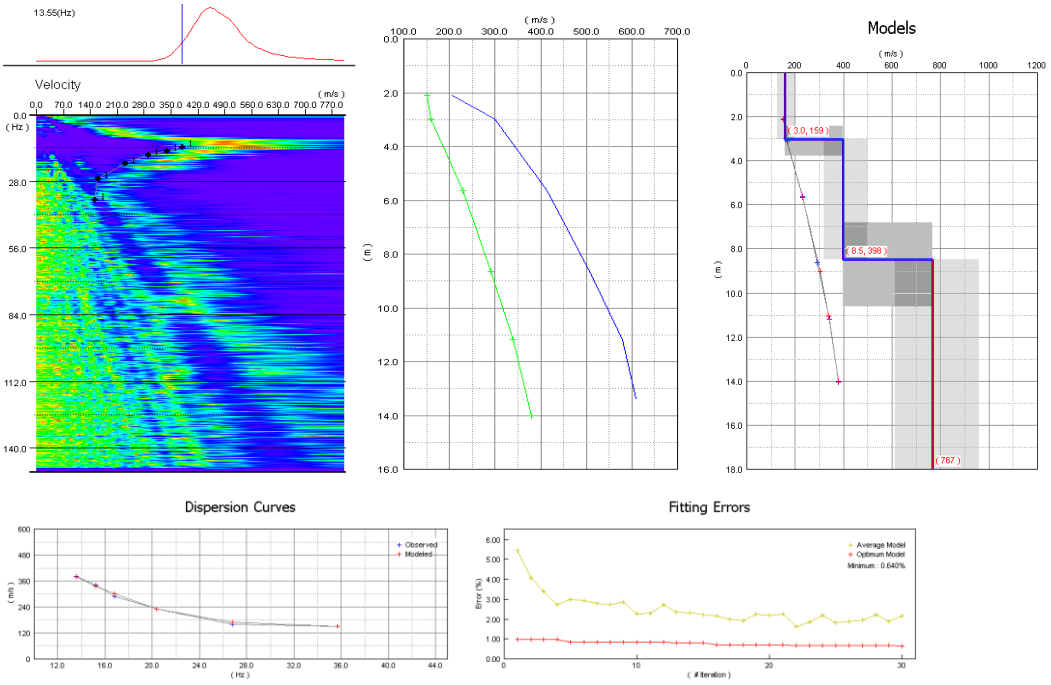
85 Meter Survey



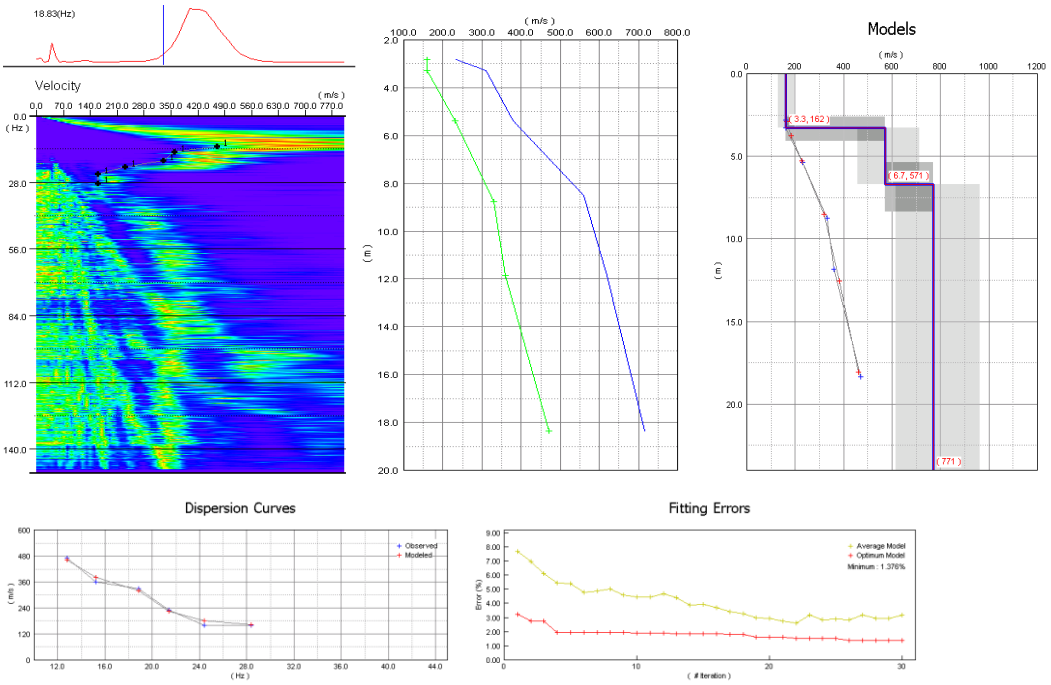
90 Meter Survey



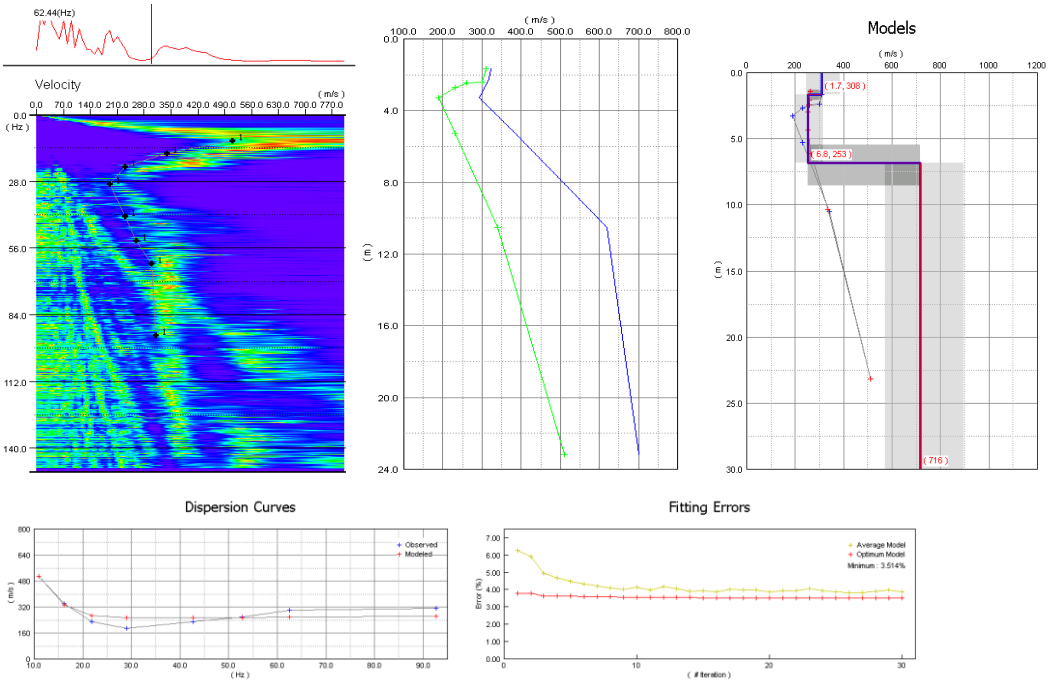
95 Meter Survey



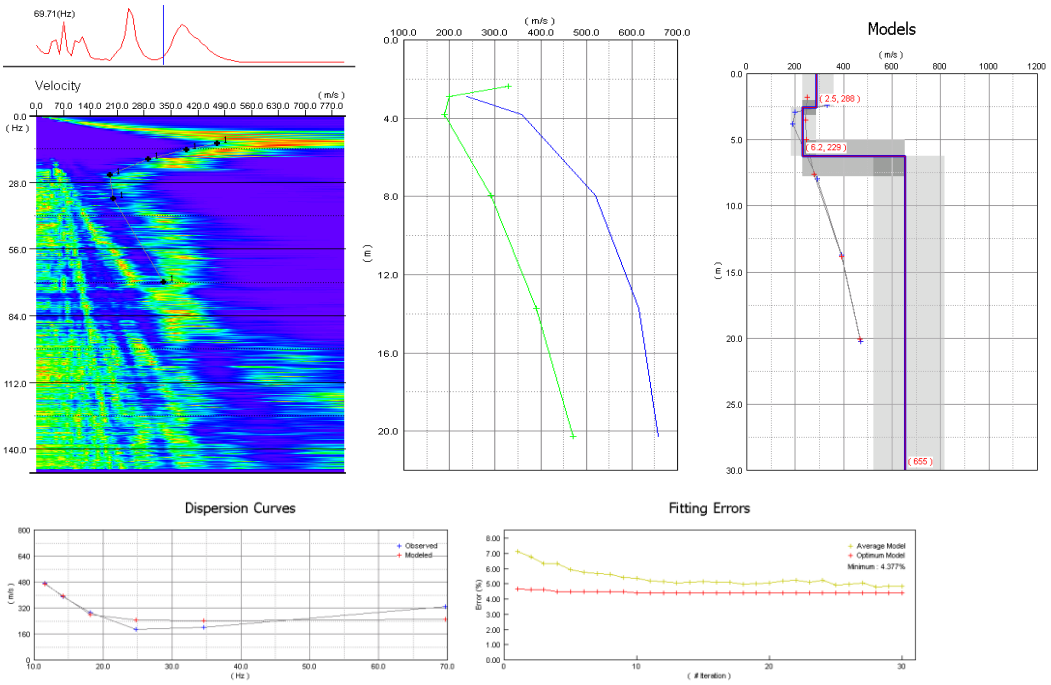
100 Meter Survey



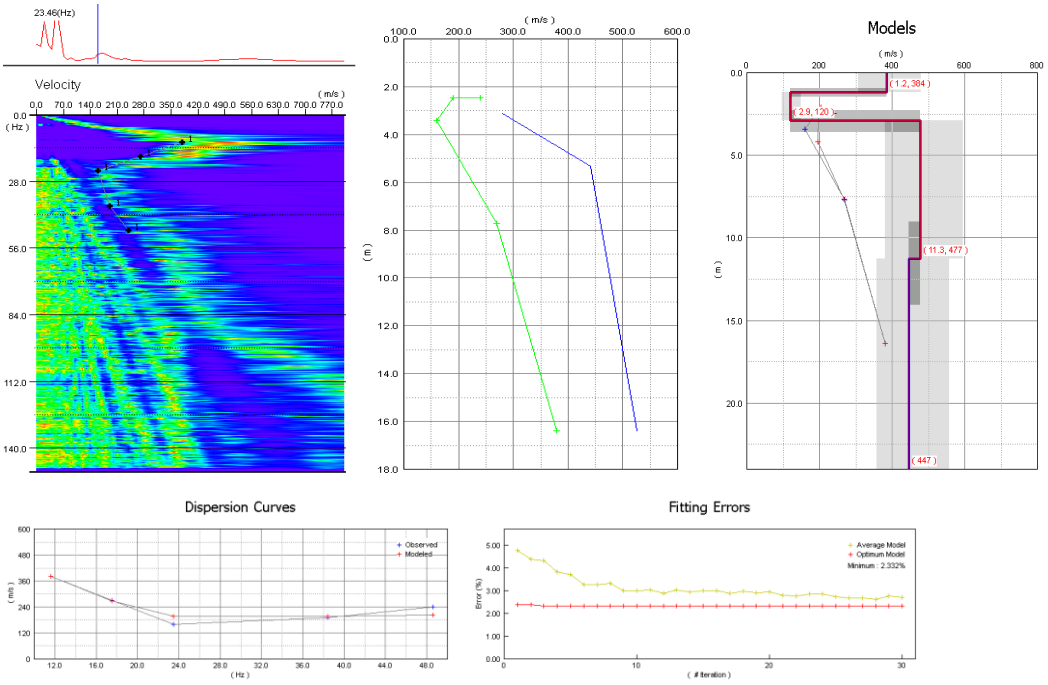
105 Meter Survey



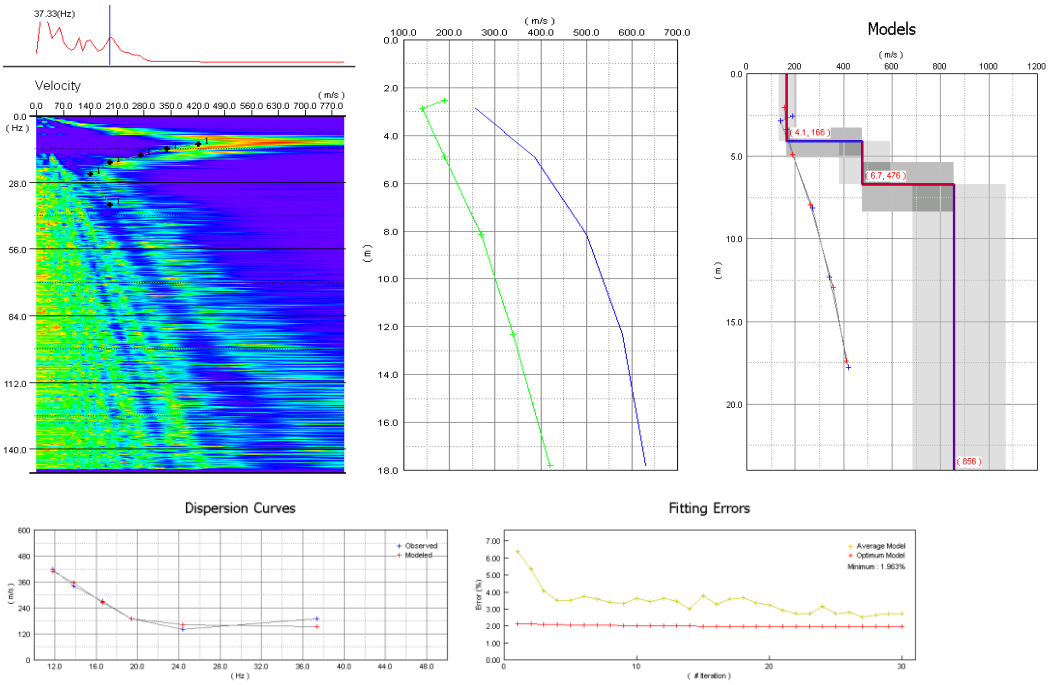
110 Meter Survey



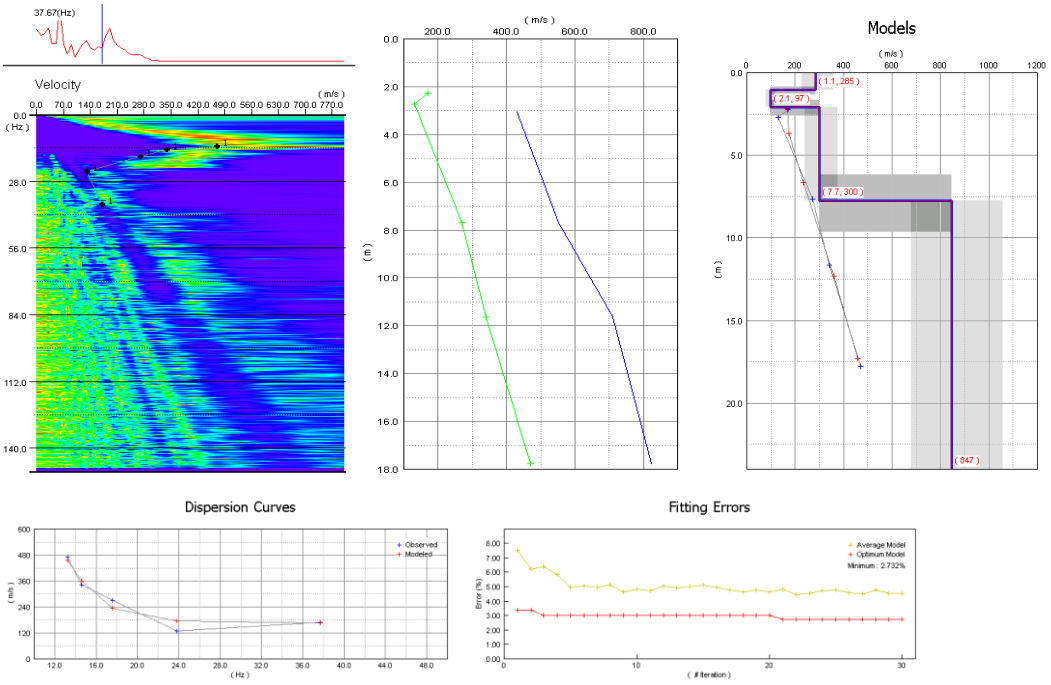
115 Meter Survey



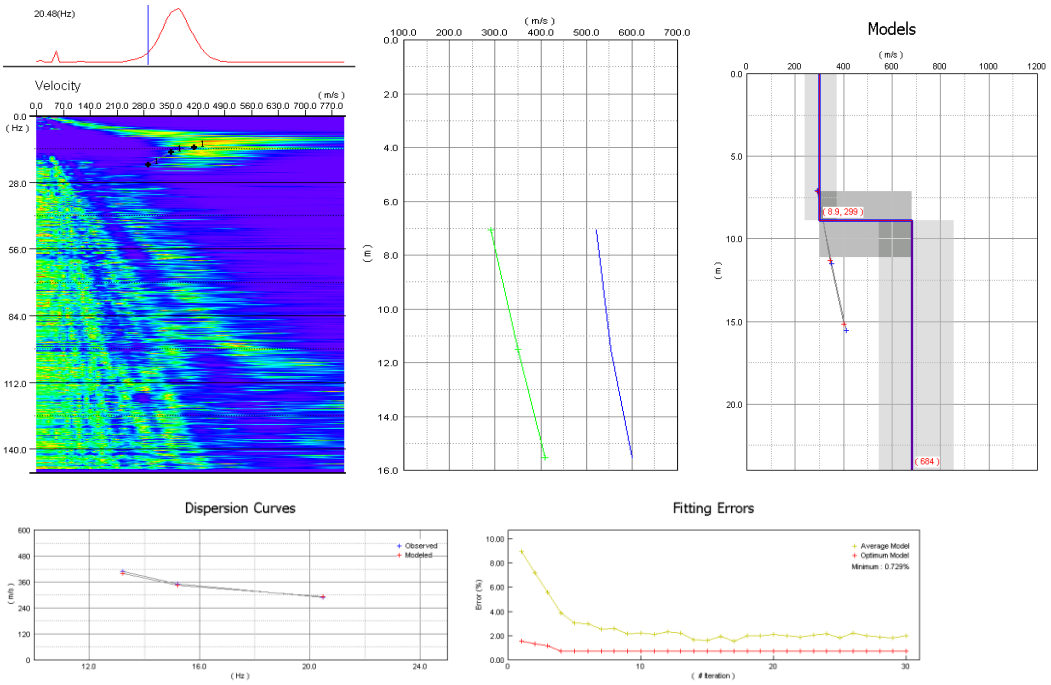
120 Meter Survey



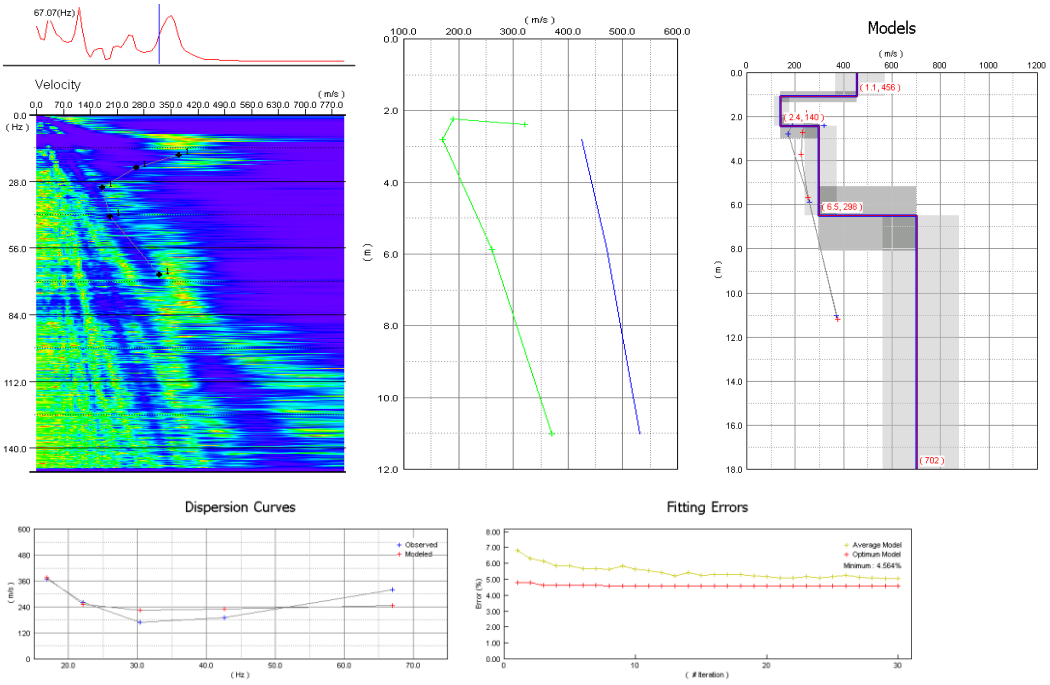
125 Meter Survey



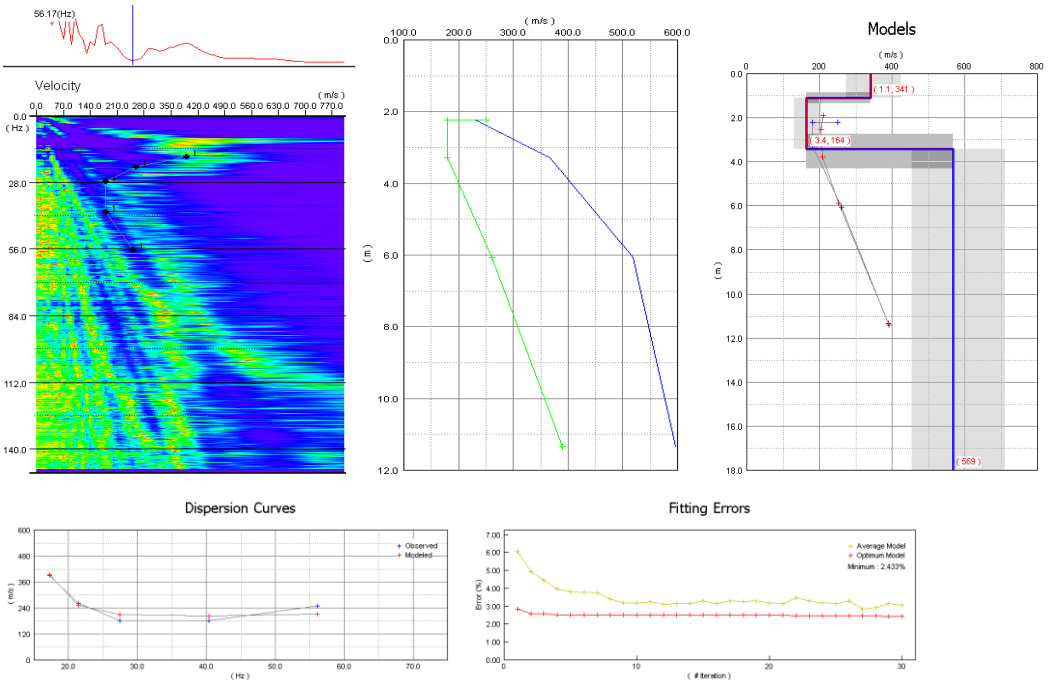
130 Meter Survey



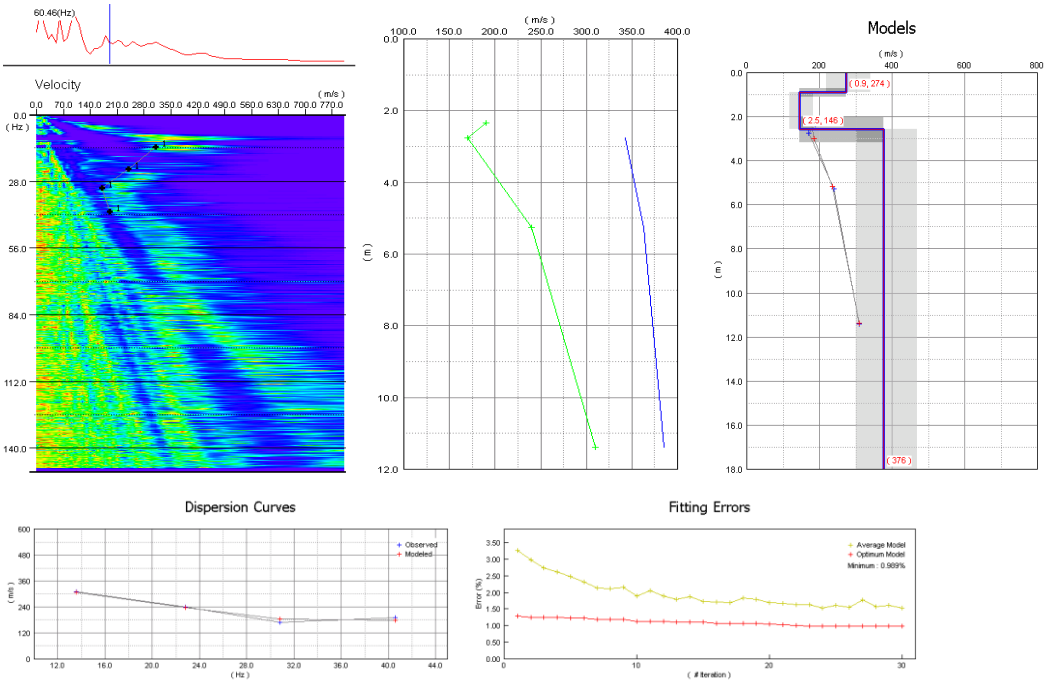
135 Meter Survey



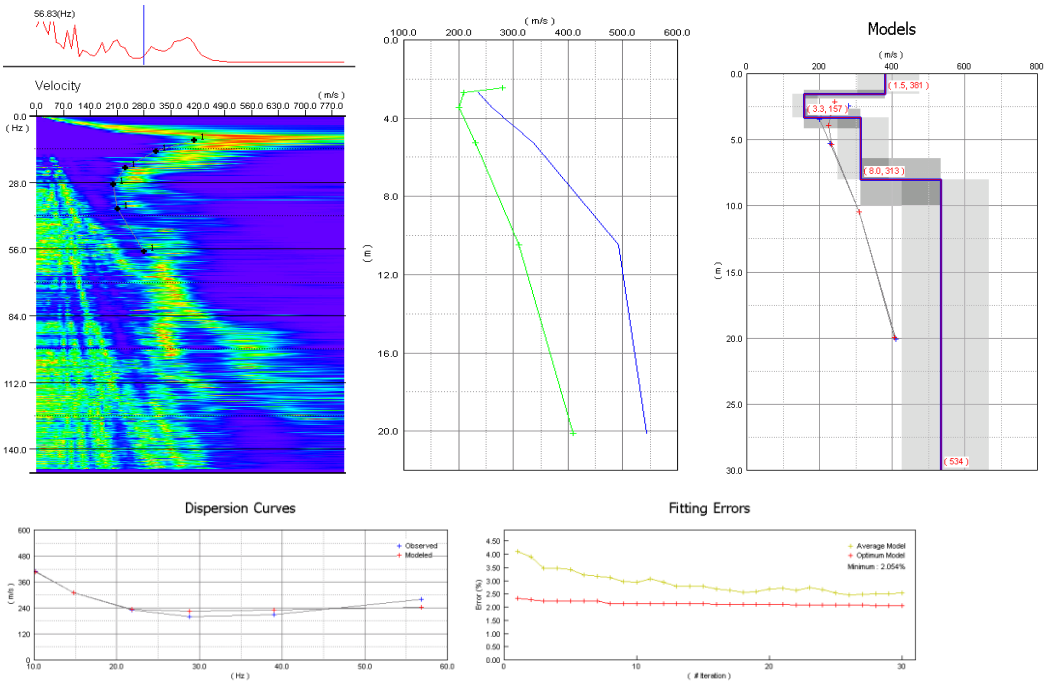
140 Meter Survey



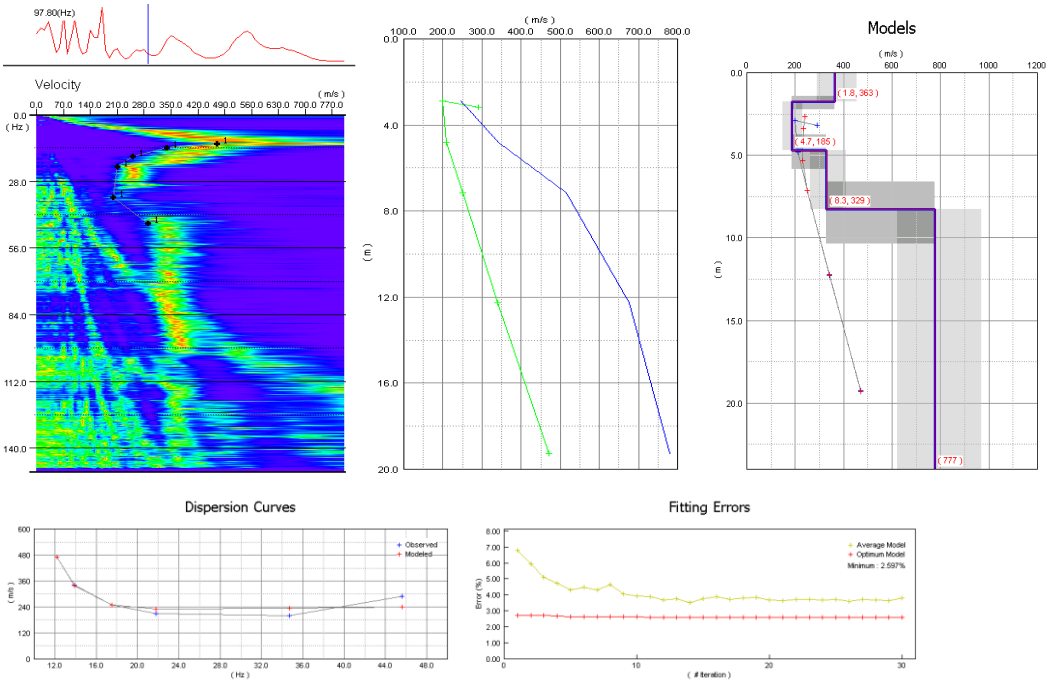
145 Meter Survey



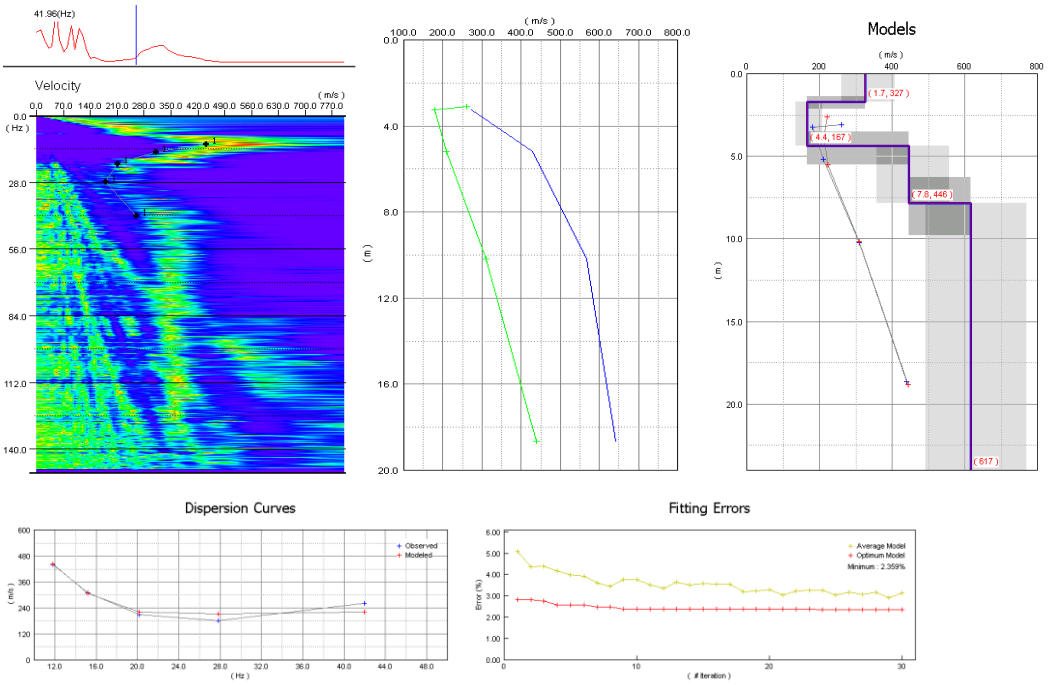
150 Meter Survey



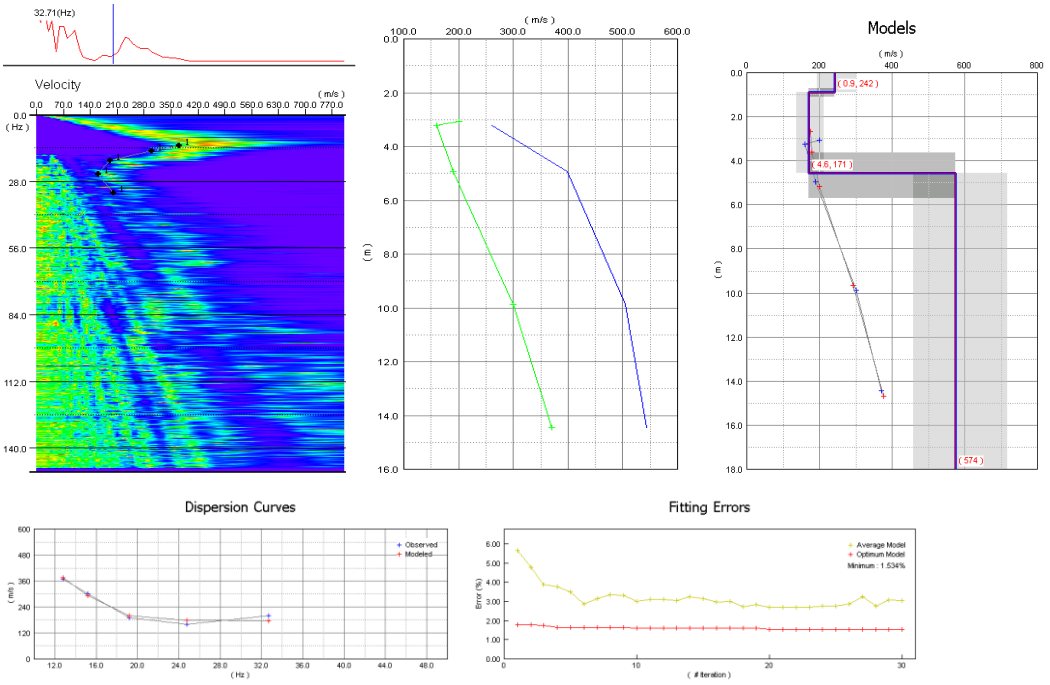
155 Meter Survey



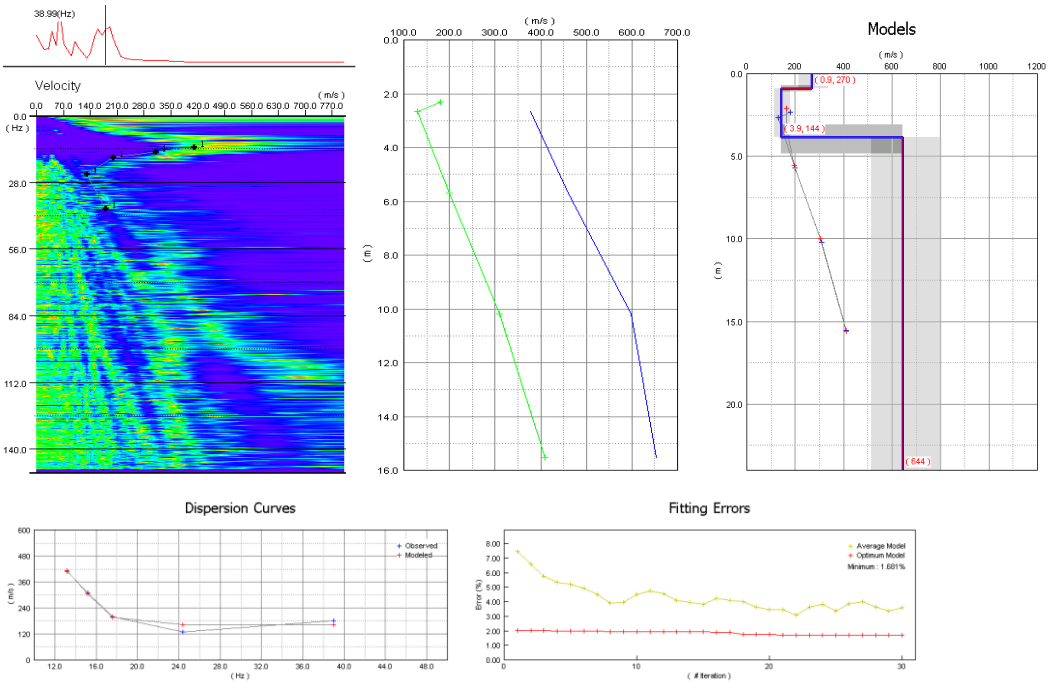
160 Meter Survey



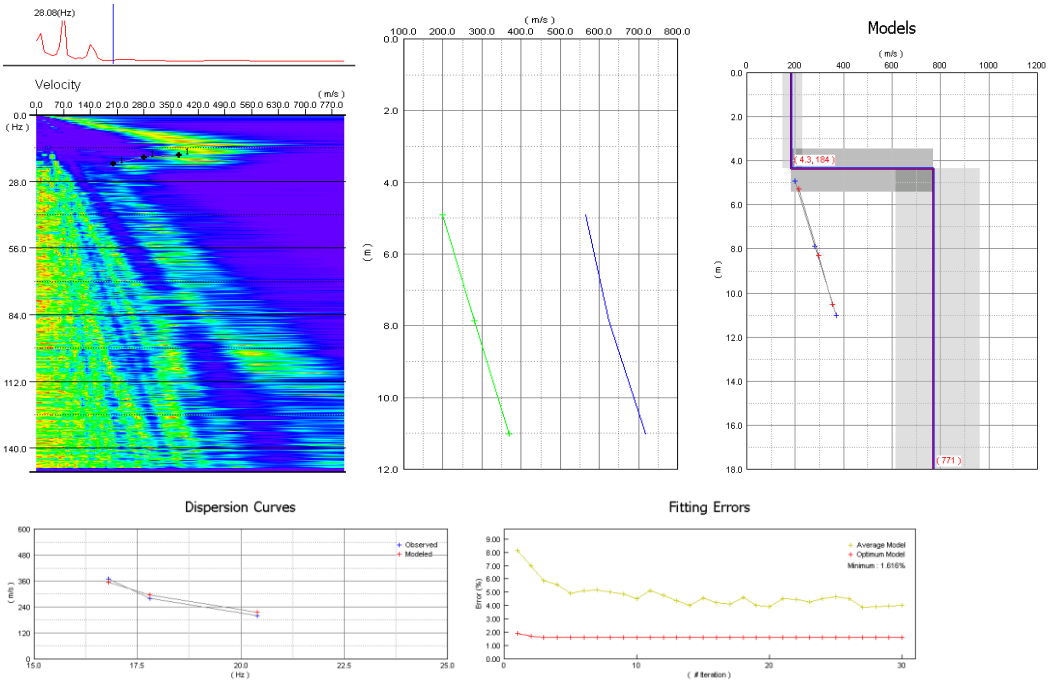
165 Meter Survey



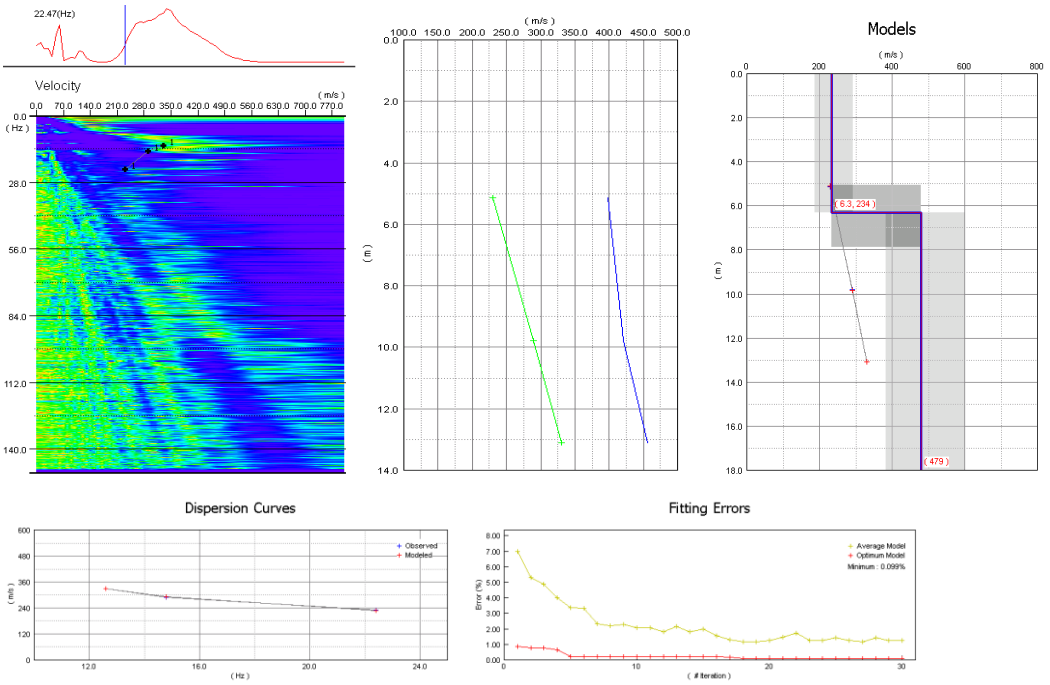
170 Meter Survey



175 Meter Survey

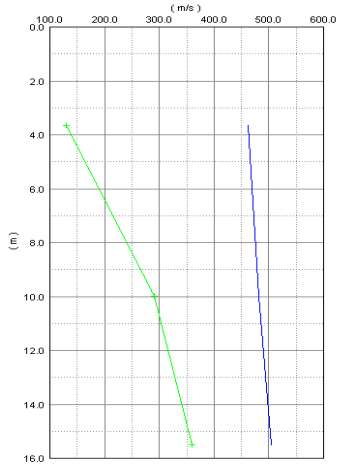


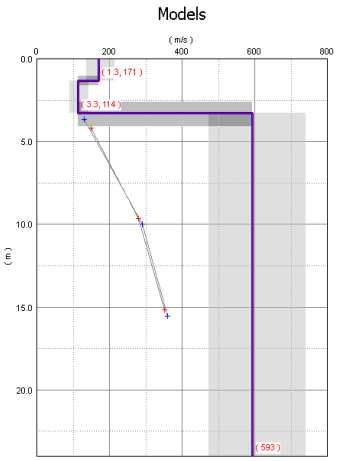
180 Meter Survey

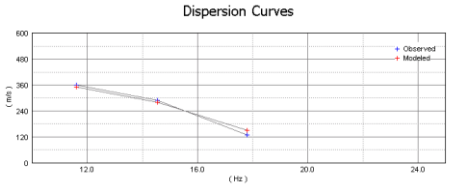


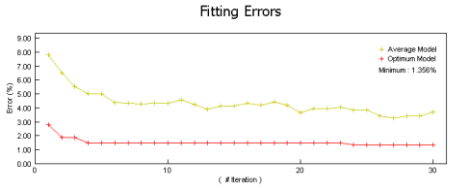
185 Meter Survey

*No V_R dispersion curve available.

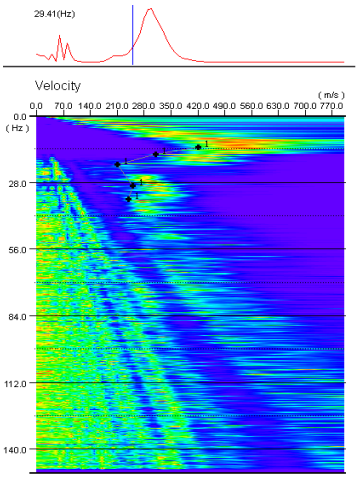


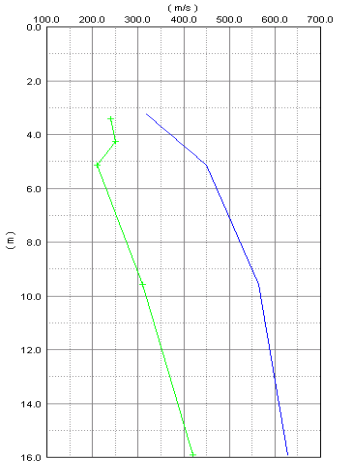


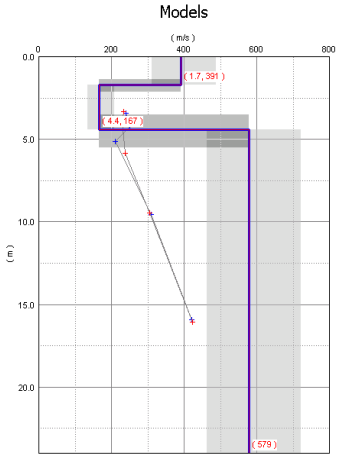


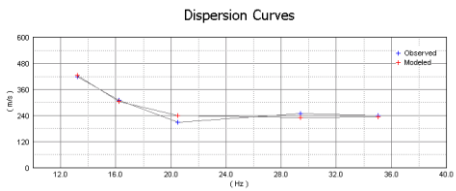


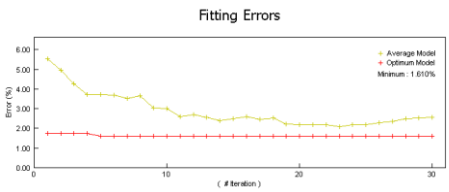
190 Meter Survey



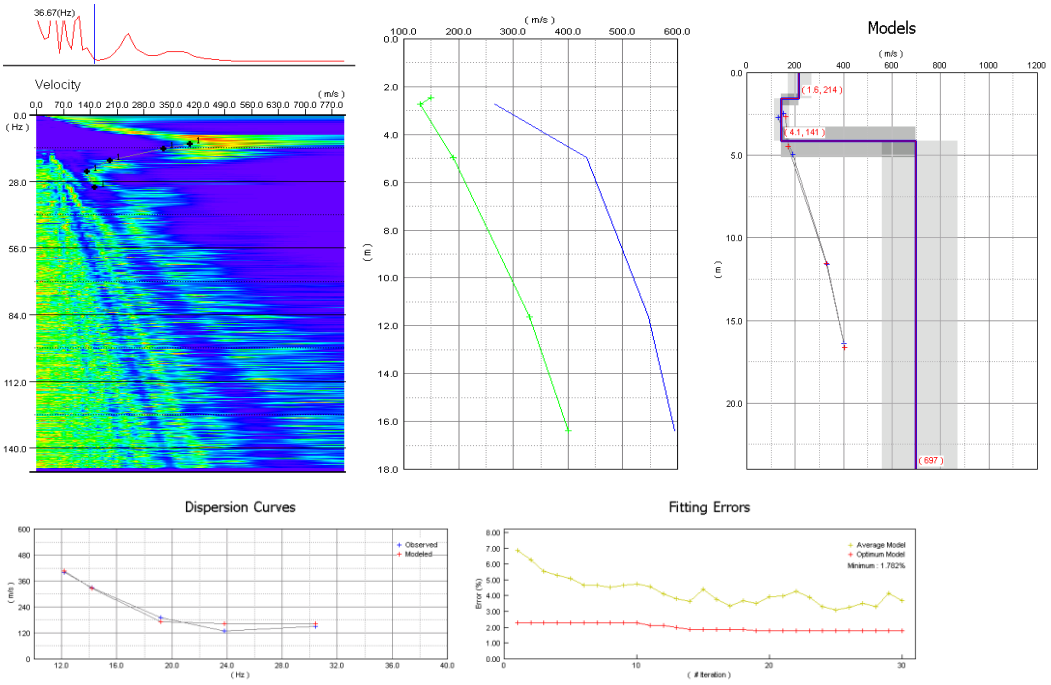




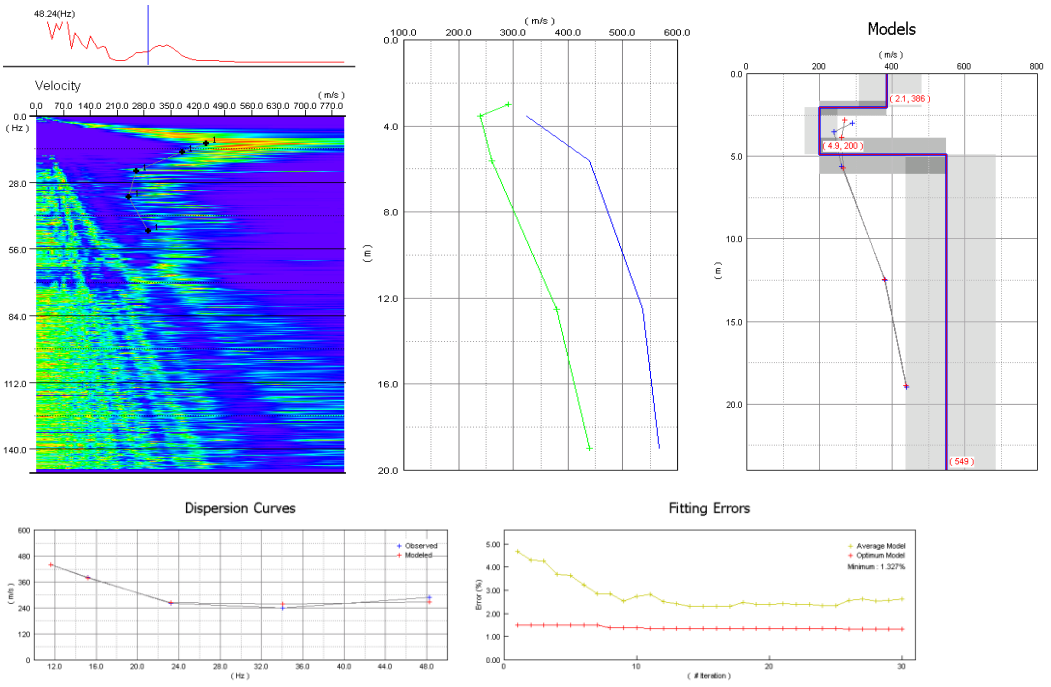




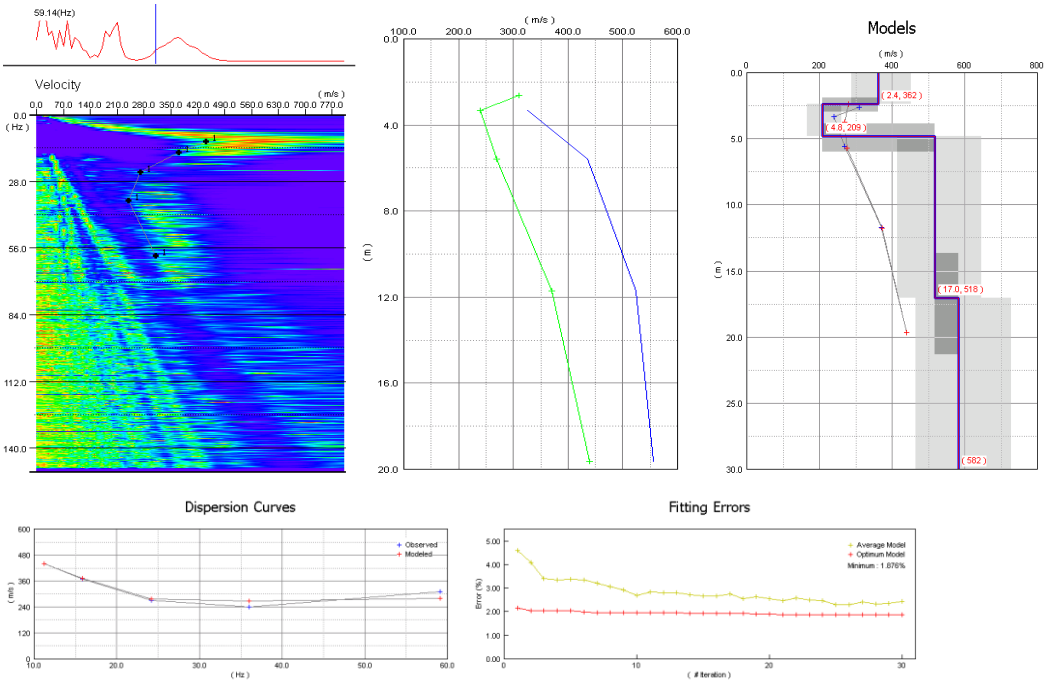
195 Meter Survey



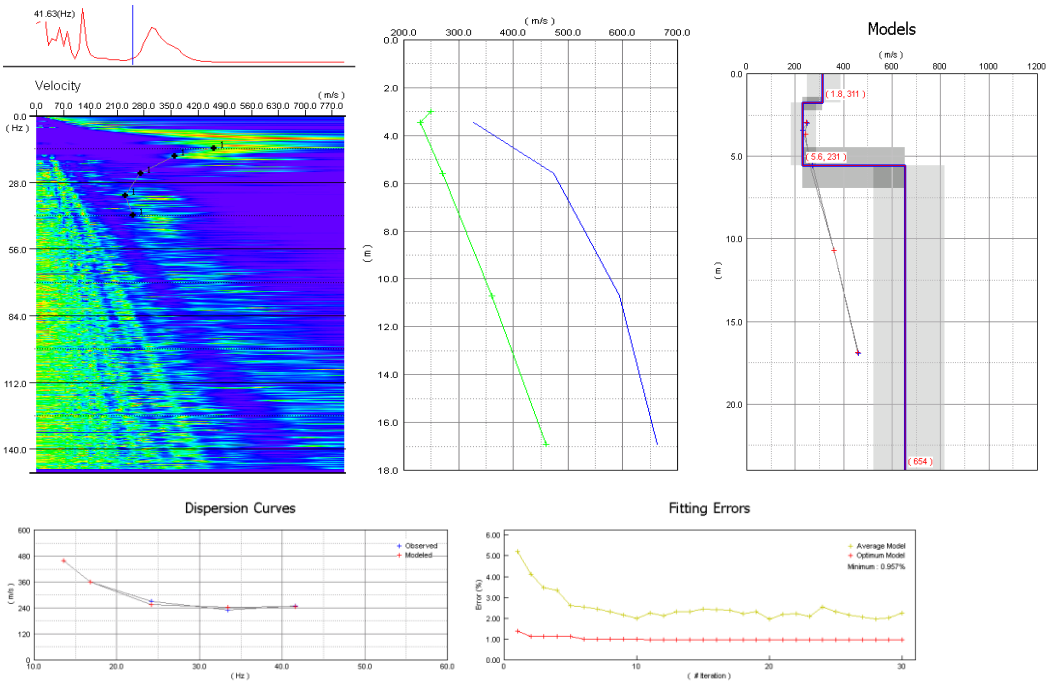
220 Meter Survey



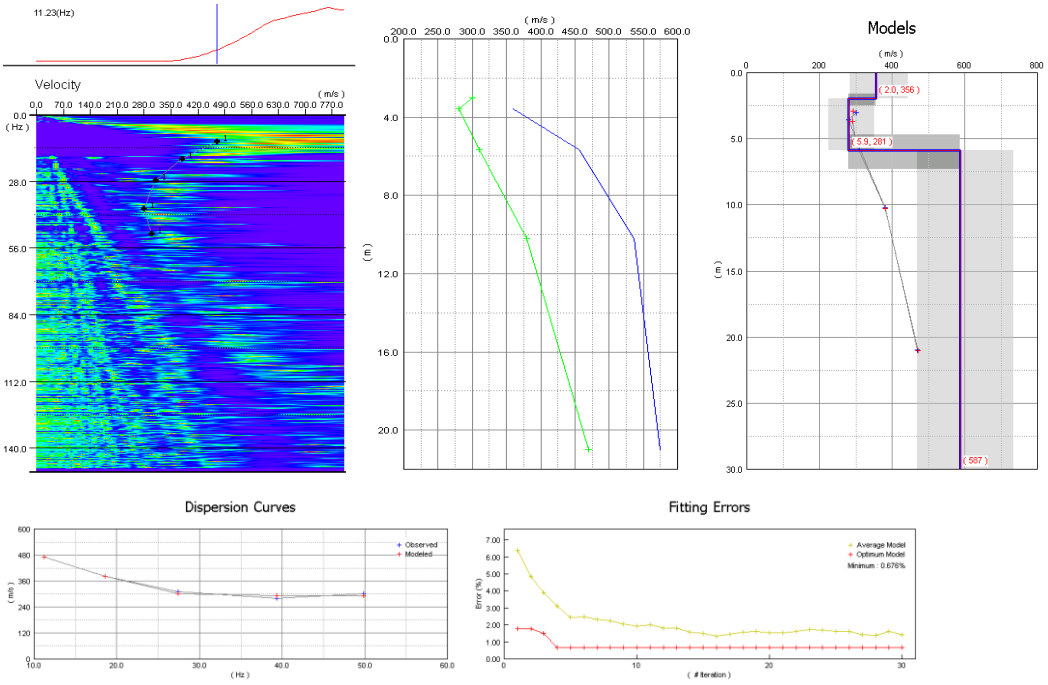
225 Meter Survey



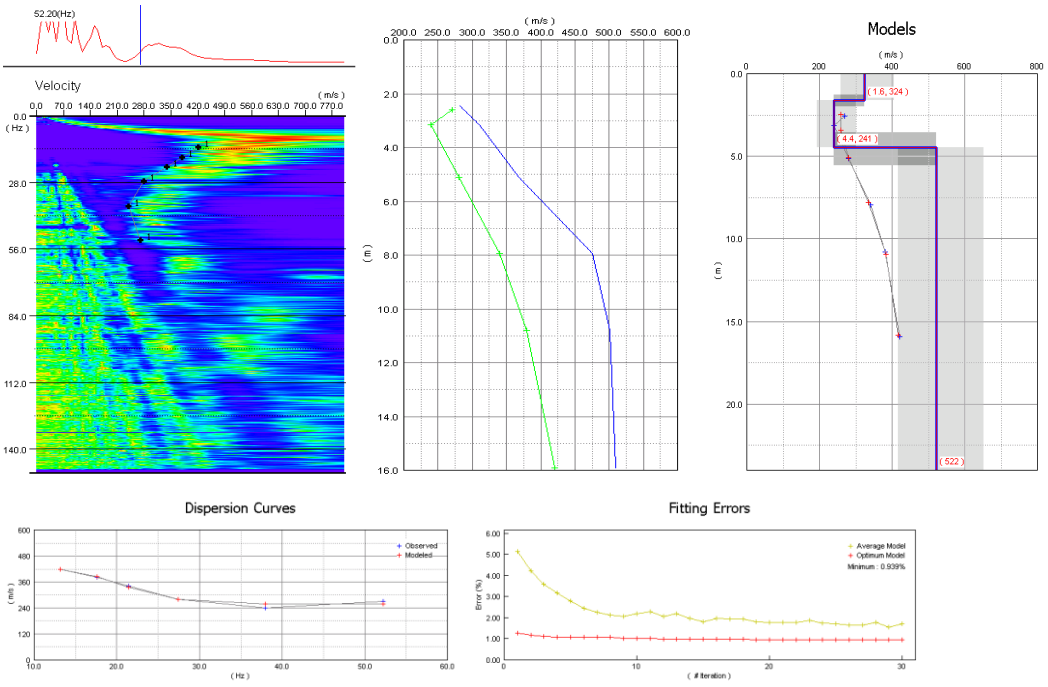
230 Meter Survey



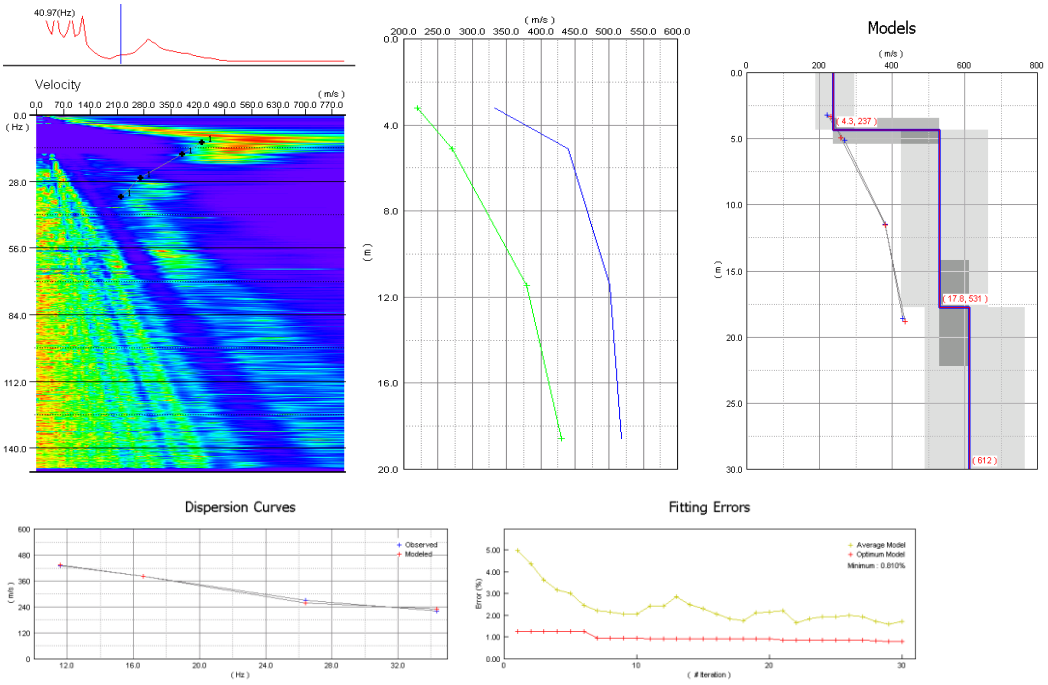
235 Meter Survey



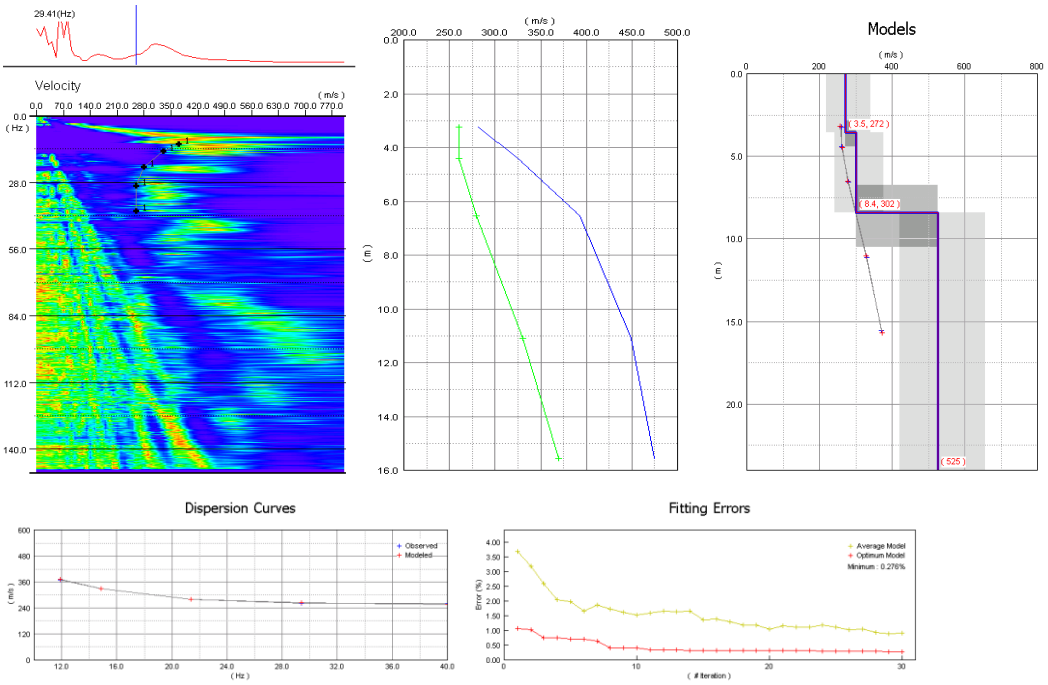
240 Meter Survey



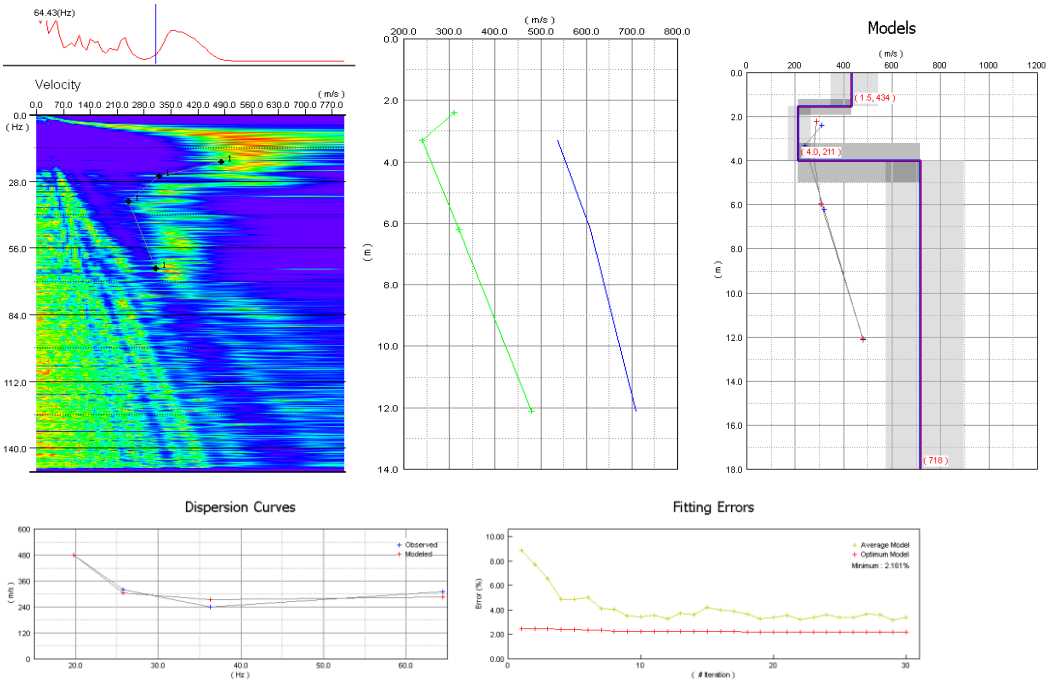
245 Meter Survey



250 Meter Survey

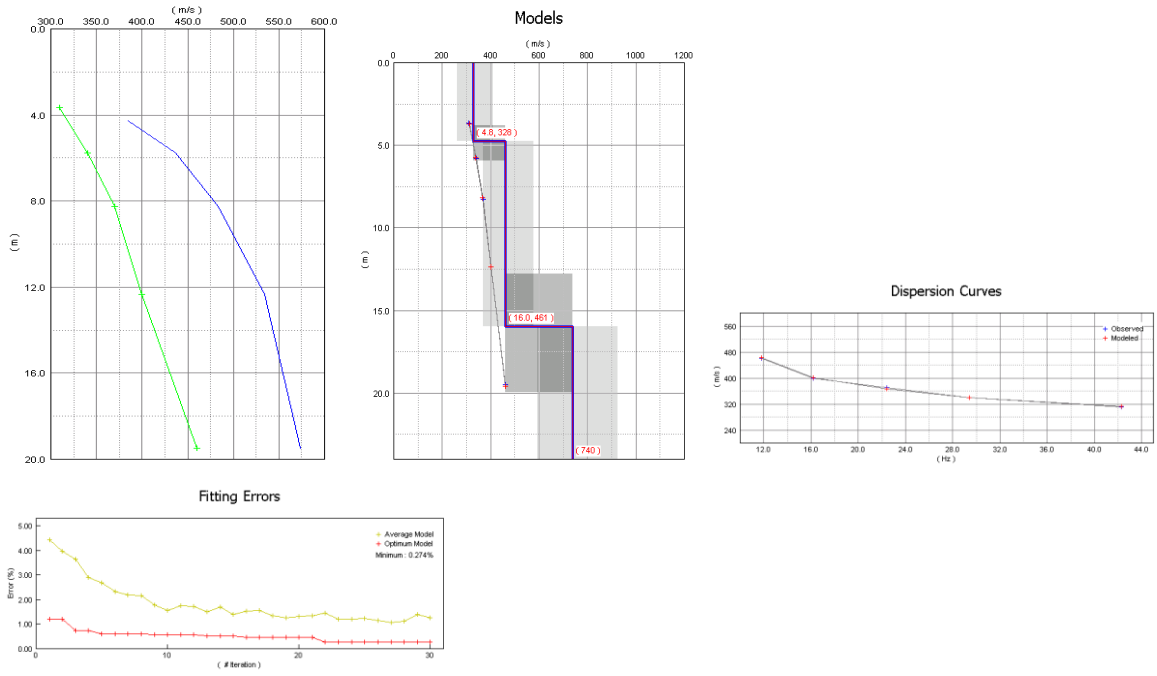


255 Meter Survey

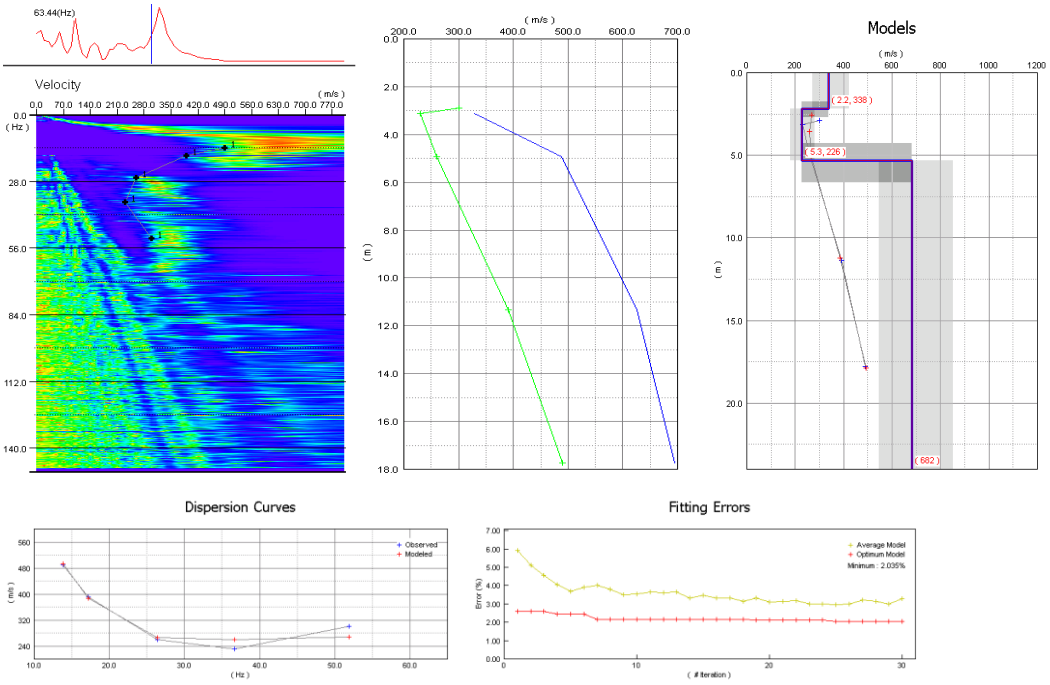


260 Meter Survey

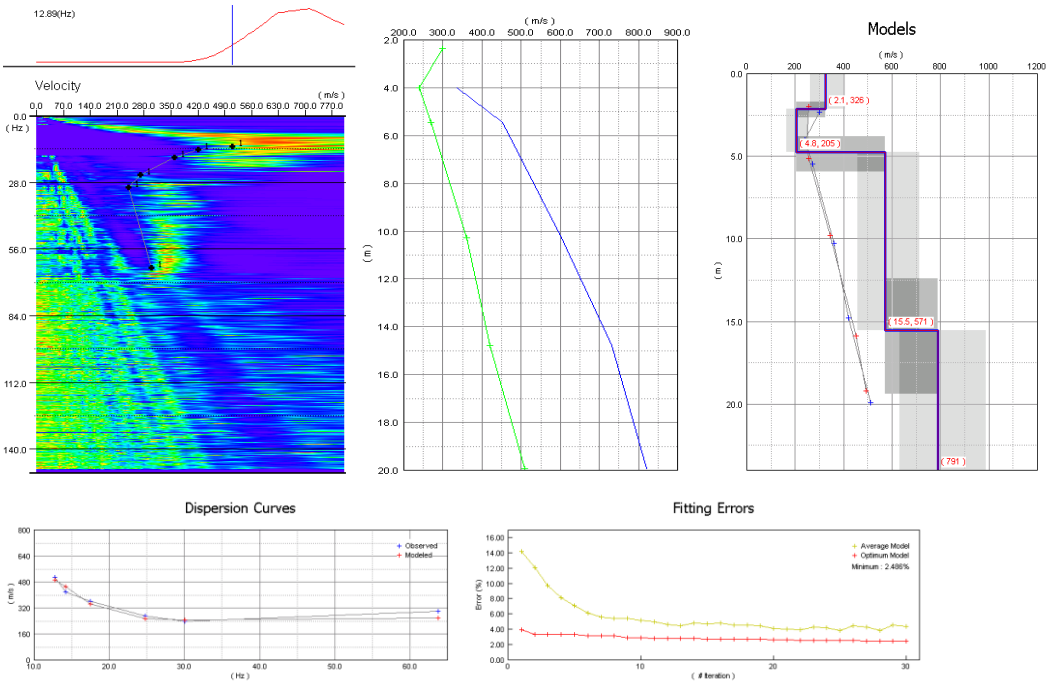
*No V_R dispersion curve available.



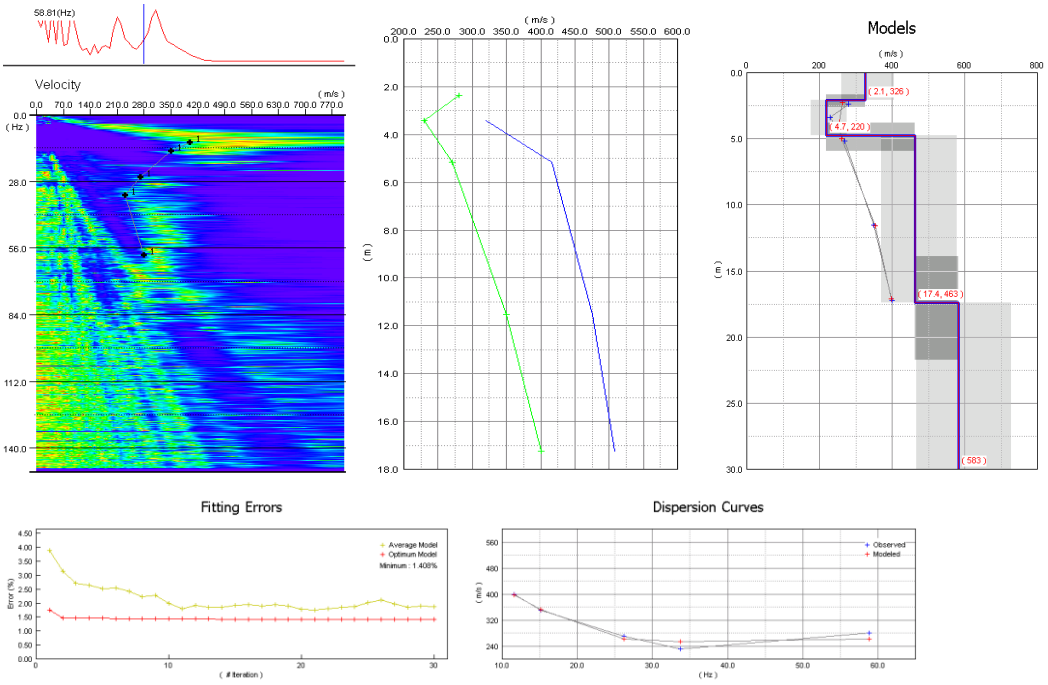
265 Meter Survey



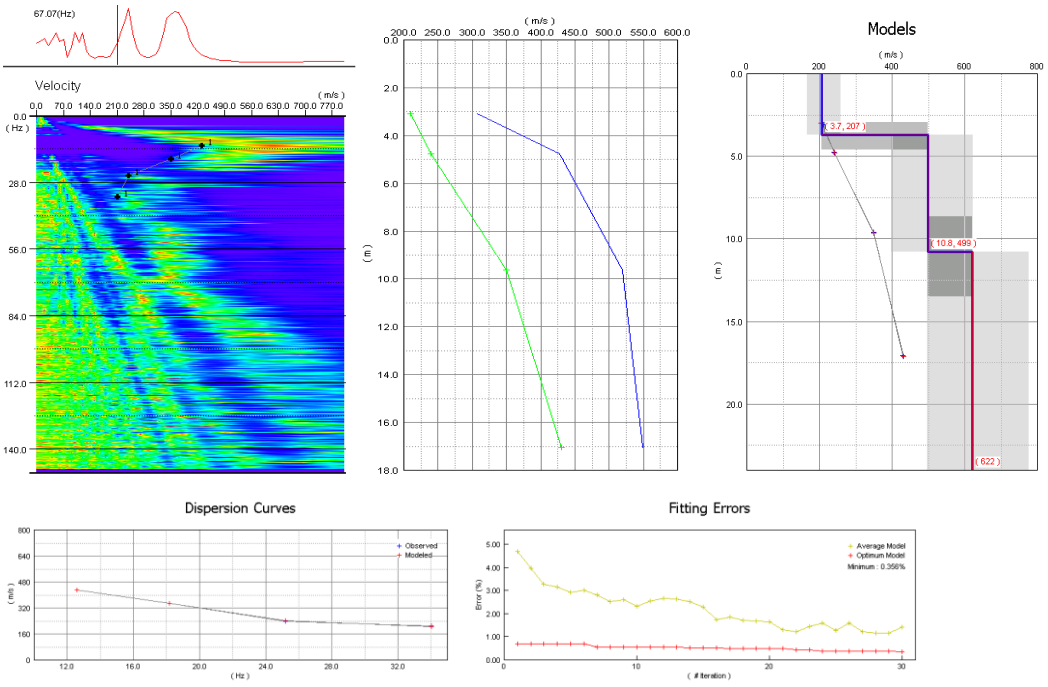
270 Meter Survey



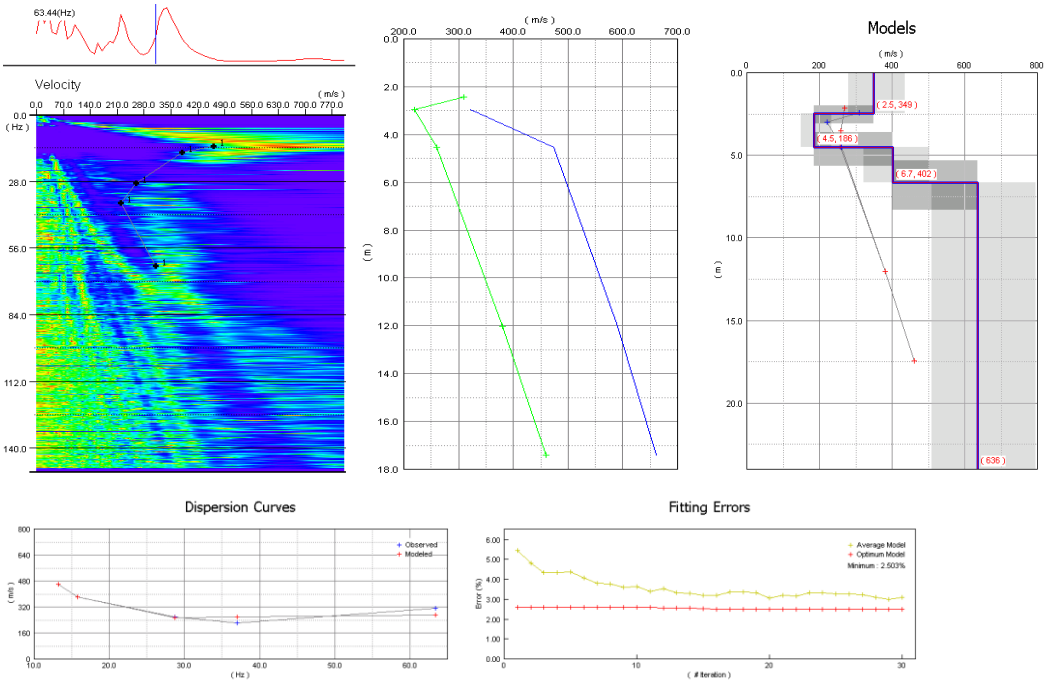
275 Meter Survey



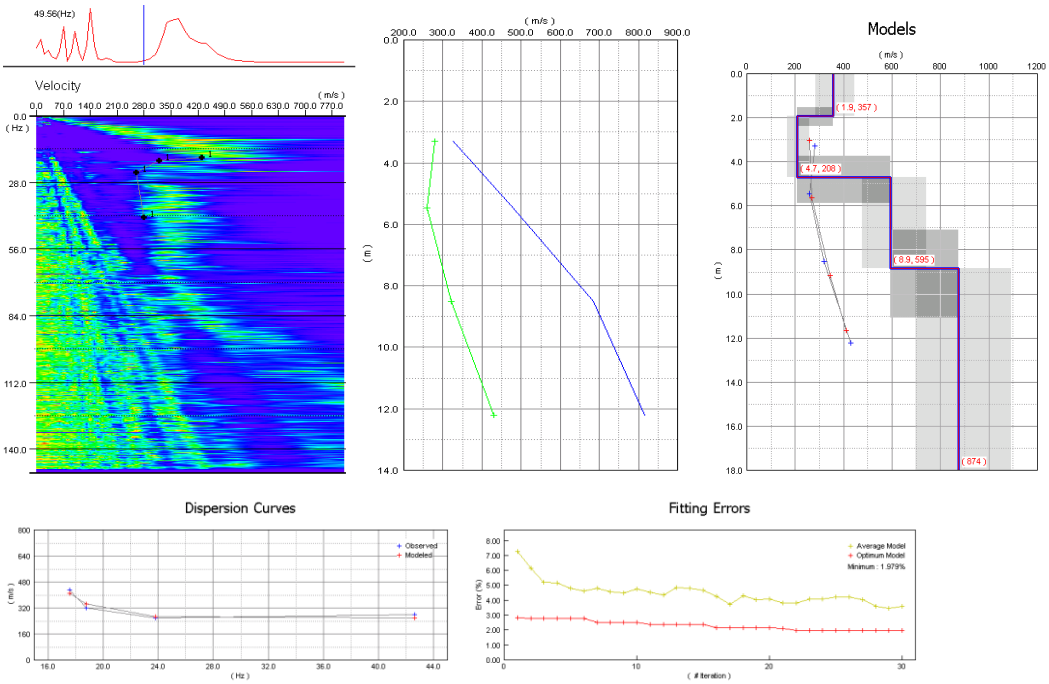
280 Meter Survey



285 Meter Survey



290 Meter Survey



Appendix C: Time weighted average Phase Velocity Values for Each Survey

The time weighted average phase velocity values for Rayleigh waves and shear waves are shown below. Note that the offset columns refer to the survey locations with respect to the reference stake (see Figure 3.2). The column headers correspond to the depths where the soil columns were analyzed.

Rayleigh wave Values:

Offset	5m (m/s)	10m (m/s)	15m (m/s)	Offset	5m (m/s)	10m (m/s)	15m (m/s)
-20	180	210	245	130		303	321
-15	174	210	250	135	225	255	
-10	179	226	268	140	218	251	
-5	182	224	254	145	195	225	
0	213	249	278	150	248	256	277
5	245	263	289	155	252	253	278
10	260	280	306	160	230	242	269
15		332	356	165	189	213	**244**
20	250	275	306	170	164	197	230
25	174	222	258	175	200	230	
30	295	309	338	180	225	243	
35	261	297	329	185	151	182	214
40	264	293	324	190	243	253	281
45	258	294	319	195	152	188	221
50	255	282	312	220	274	294	318
55	263	280	307	225	277	299	321
60	176	208	239	230	257	286	315
65	284	313	337	235	279	310	335
70		305	326	240	271	302	**328**
75	222	262	297	245	241	275	303
80	260	284	308	250	265	283	**302**
85		290		255	293	332	
90	213	243	268	260	317	341	360
95	163	202	236	265	279	305	335
100	164	209	245	270	278	302	329
105	247	265	292	275	267	289	310
110	243	258	286	280	229	265	295
115	210	231	257	285	273	298	325
120	180	210	242	290	273	299	
125	169	210	246				

Key:		
	Outlier=	Value
	Extrapolated=	**Value**
	Not Available=	

Shear wave Velocity:

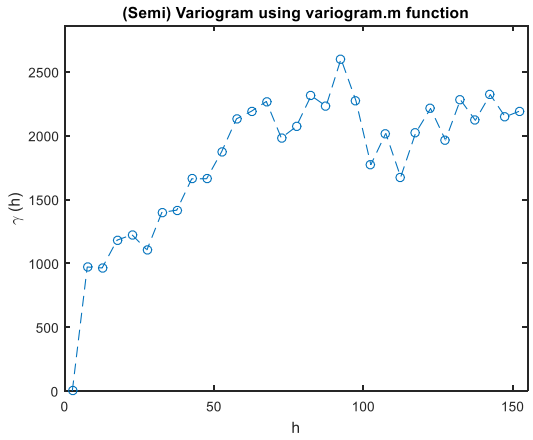
Offset	5m (m/s)	10m (m/s)	15m (m/s)	Offset	5m (m/s)	10m (m/s)	15m (m/s)
-20	193	302	384	130		319	388
-15	211	334	422	135	245	328	
-10	236	337	429	140	249	346	
-5	234	350	360	145	239	292	
0	255	312	367	150	240	293	345
5	259	310	383	155	232	296	373
10	259	334	377	160	220	307	369
15		376	424	165	192	288	345
20	248	345	395	170	193	297	362
25	240	329	378	175	206	325	
30	293	368	418	180	234	289	
35	295	385	429	185	178	274	334
40	298	346	415	190	232	331	386
45	293	375	415	195	189	297	367
50	266	363	412	220	301	371	410
55	262	343	399	225	313	375	408
60	190	283	354	230	287	375	429
65	324	368	418	235	299	401	442
70		353	407	240	325	386	**418**
75	254	350	417	245	301	368	404
80	268	355	391	250	286	334	**374**
85		328		255	359	454	
90	238	298	346	260	361	397	422
95	209	289	**364**	265	308	399	455
100	214	326	404	270	298	372	422
105	269	332	404	275	301	353	379
110	255	326	392	280	286	353	404
115	230	311	347	285	293	378	429
120	188	294	376	290	309	408	
125	209	281	362				

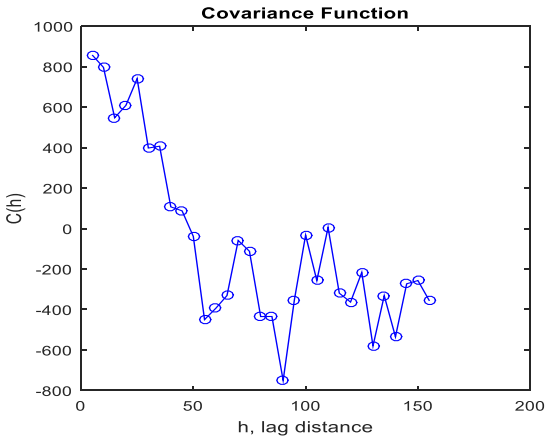
Key:		
	Outlier=	Value
	Extrapolated=	**Value**
	Not Available=	

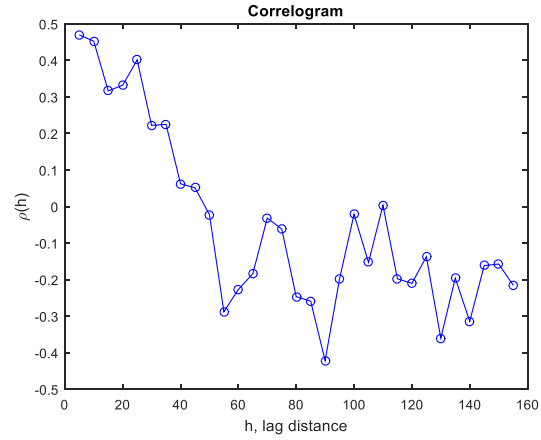
Appendix D: MATALB Output: Semi-variograms, Covariance Functions, and Correlograms

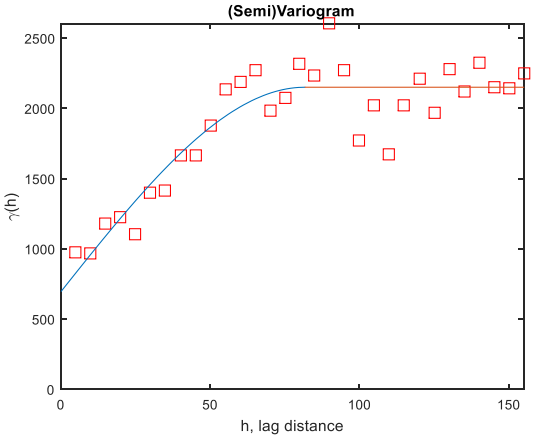
The following section contains empirical semi-variograms, covariance functions, correlograms, semi-variograms fitted with spherical models, semi-variograms fitted with exponential models, and semi-variograms fitted with Gaussian models for each depth and each differing amount of data. Note that the modeled semi-variograms are shown in the order of the spherical model, then exponential model, and then Gaussian model.

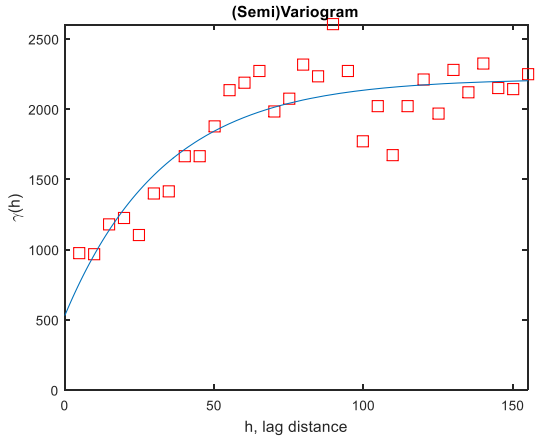
5m Rayleigh wave Velocity with All Data

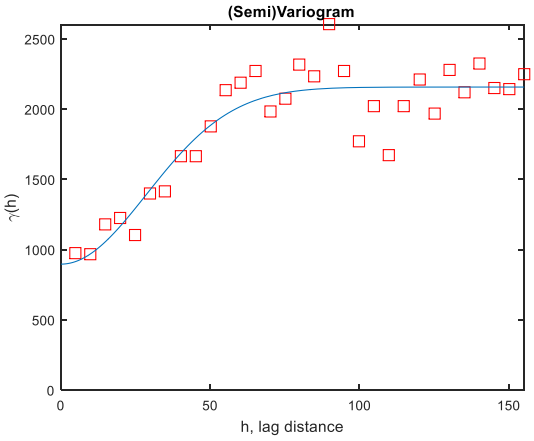




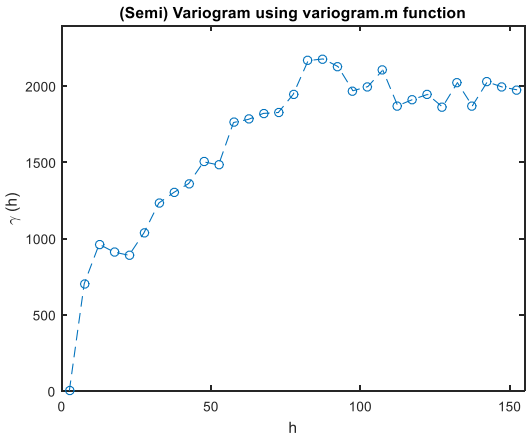


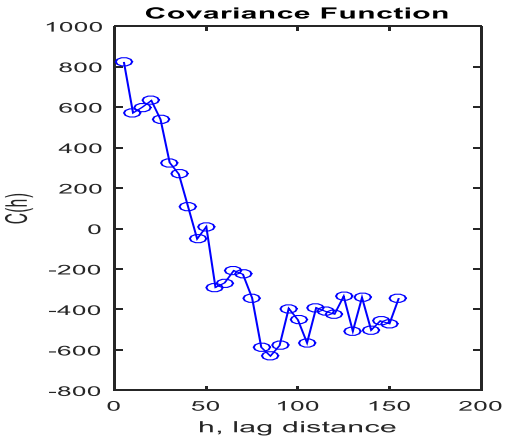


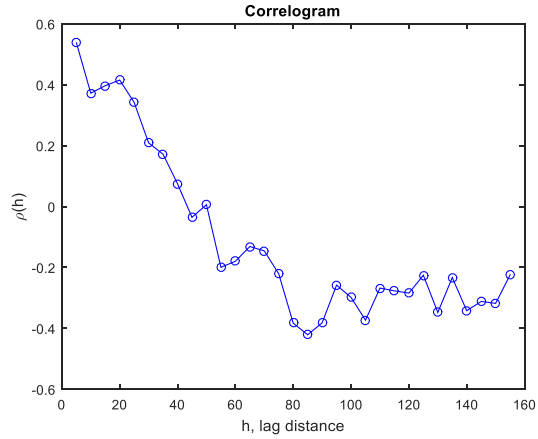


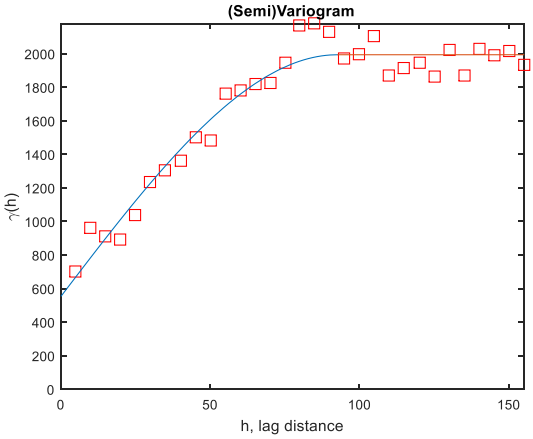


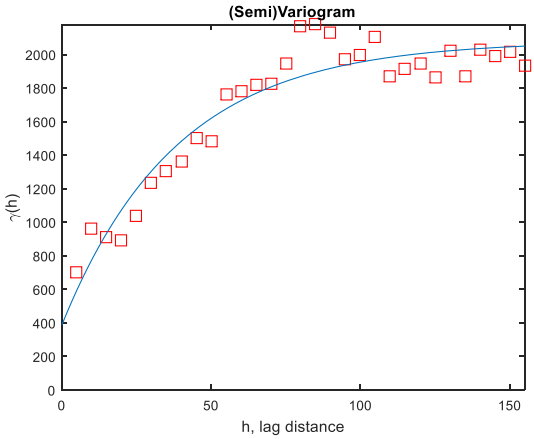
10m Rayleigh wave Velocity with All Data

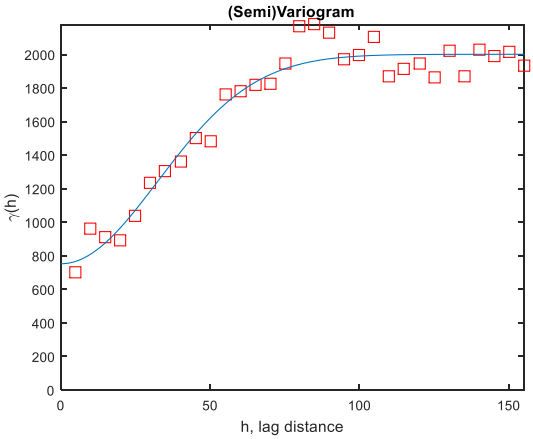




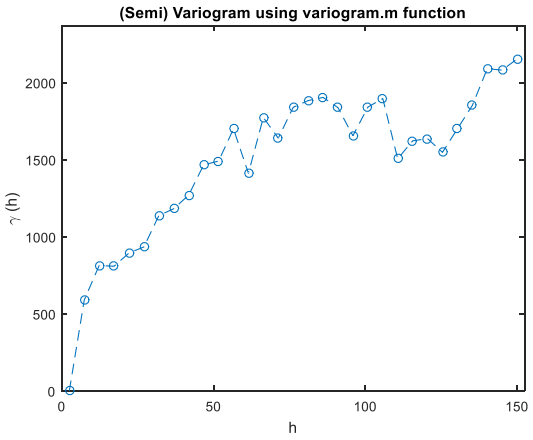


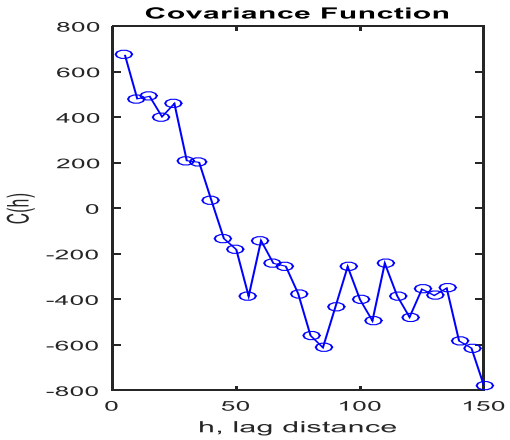


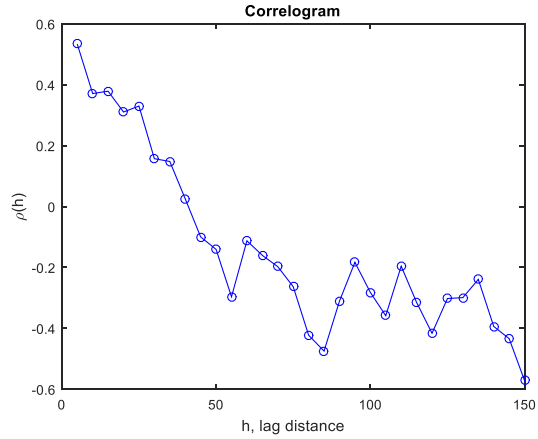


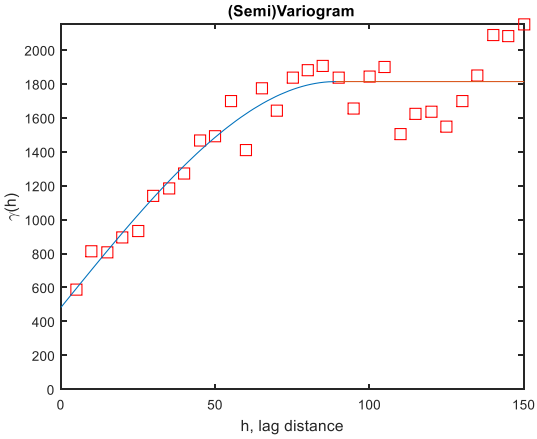


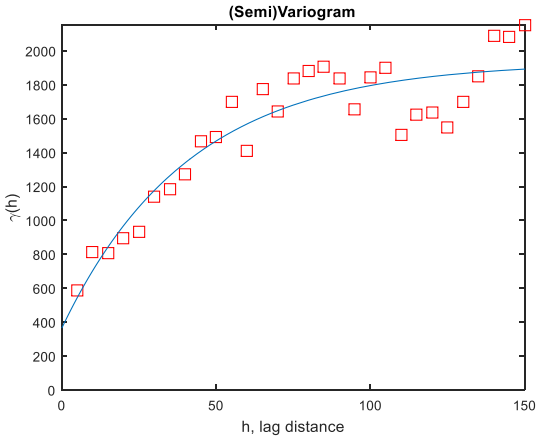
15m Rayleigh wave Velocity with All Data

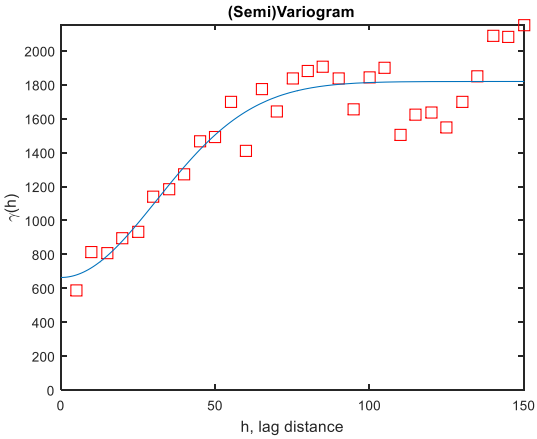




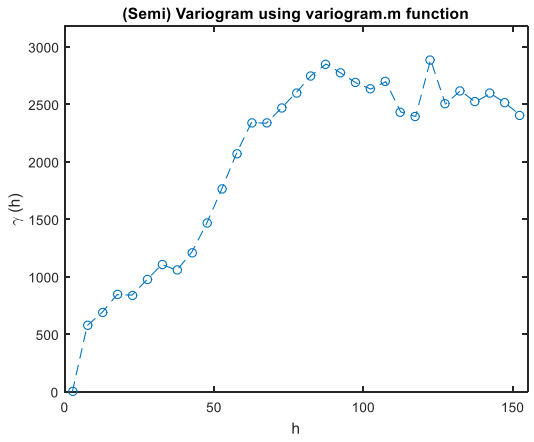


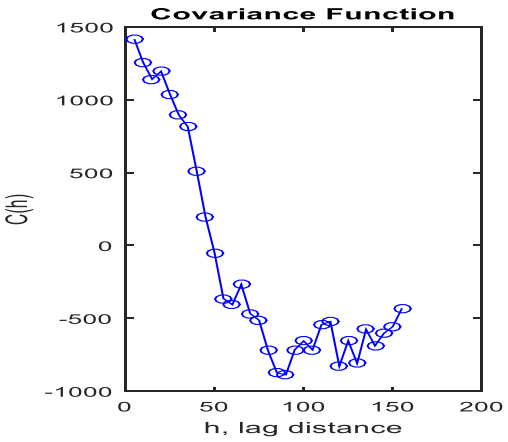


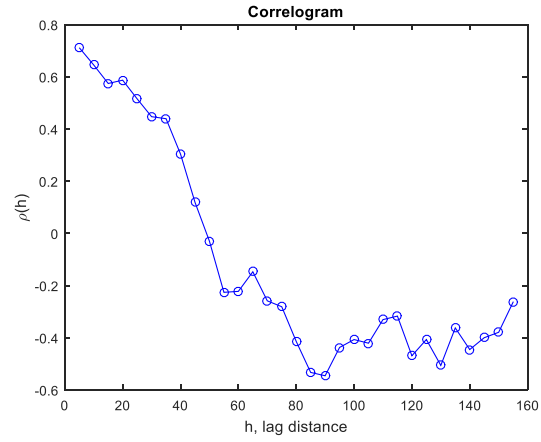


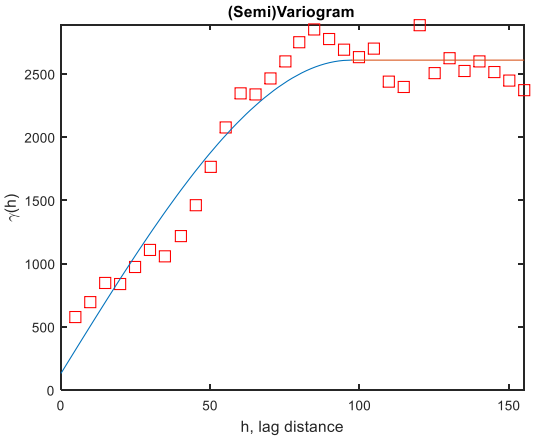


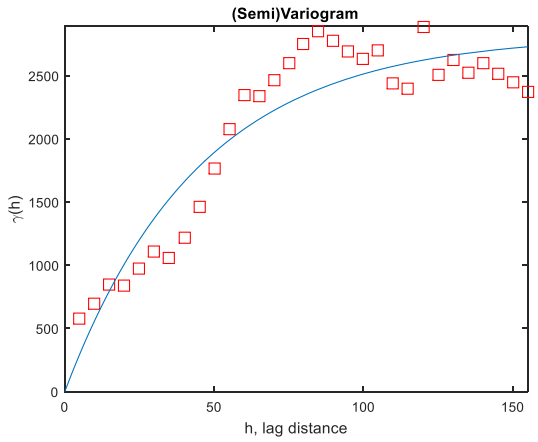
5m Shear wave Velocity with All Data

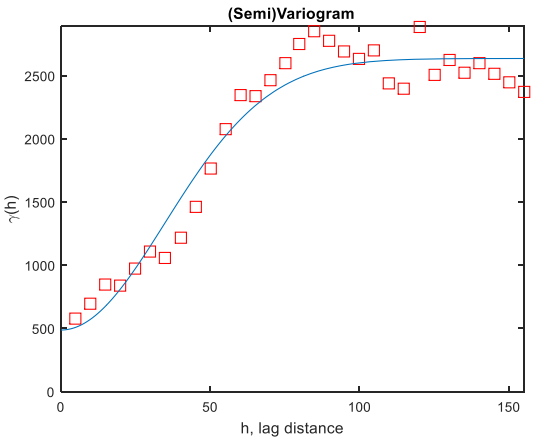




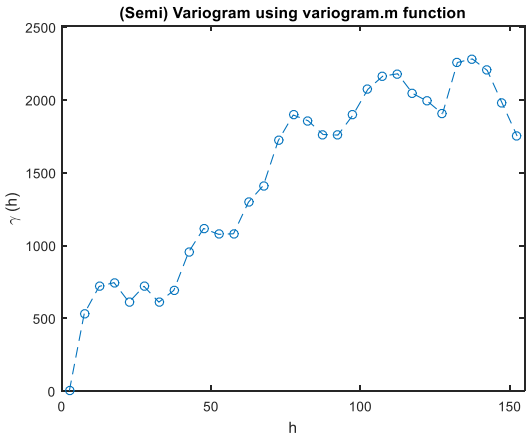


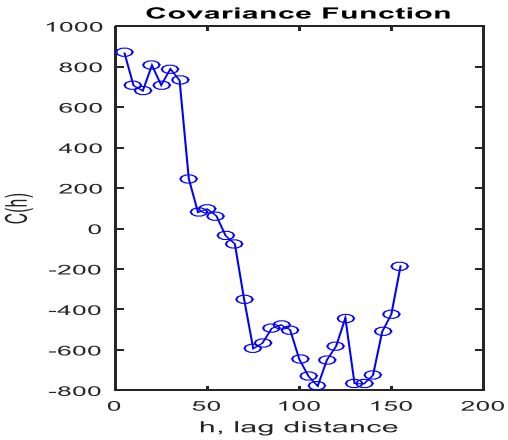


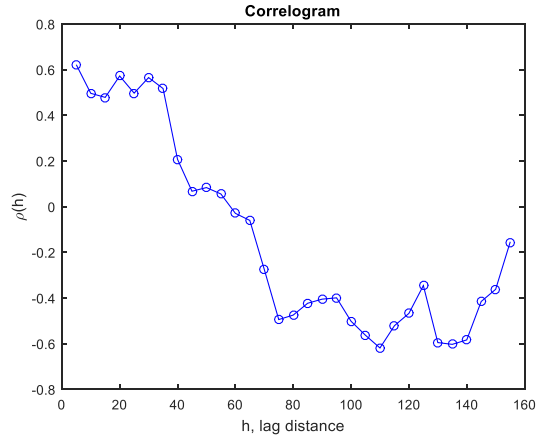


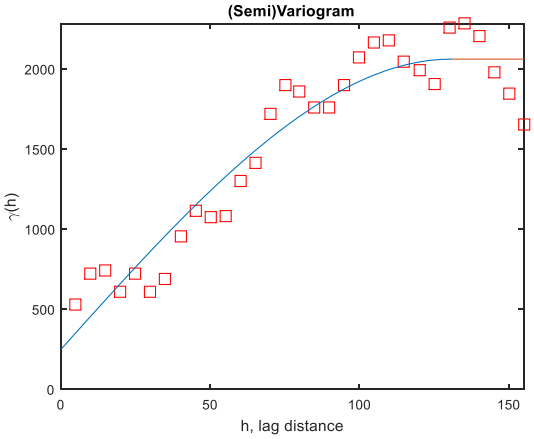


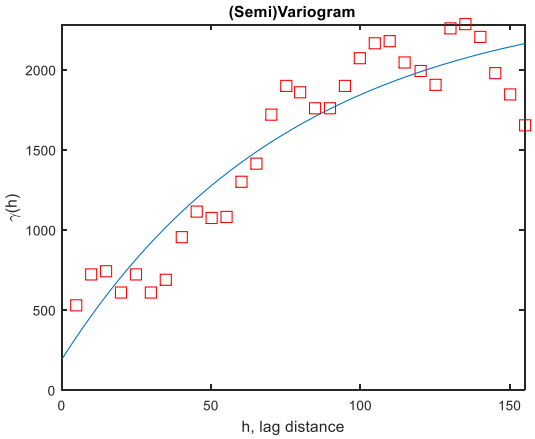
10m Shear wave Velocity with All Data

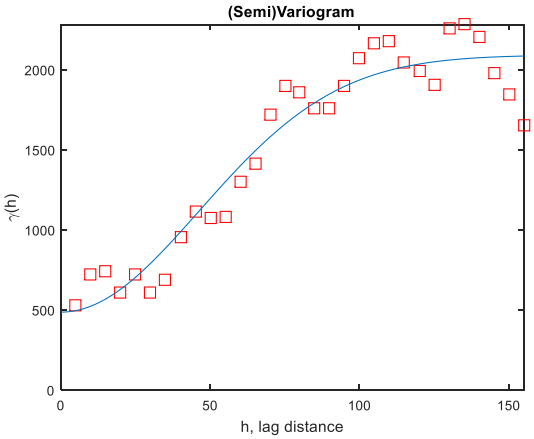




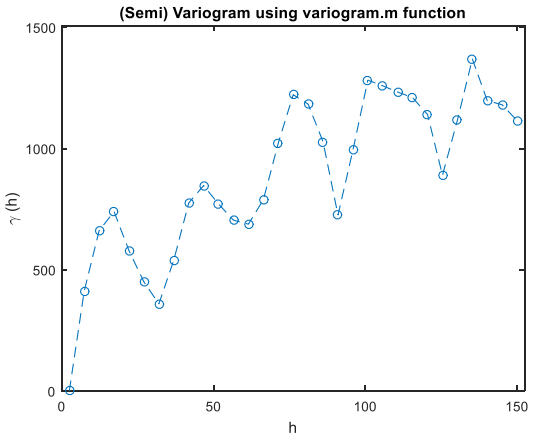


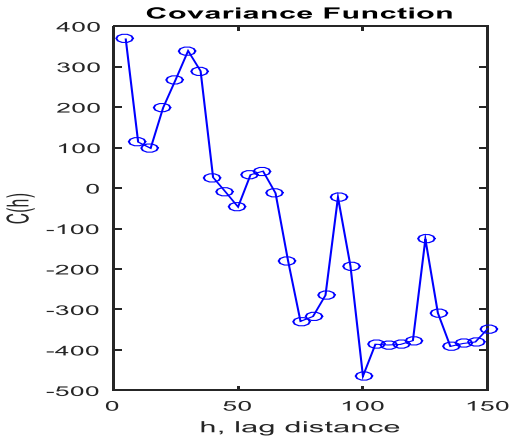


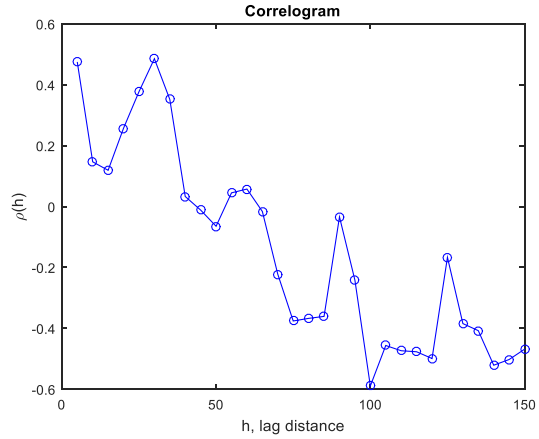


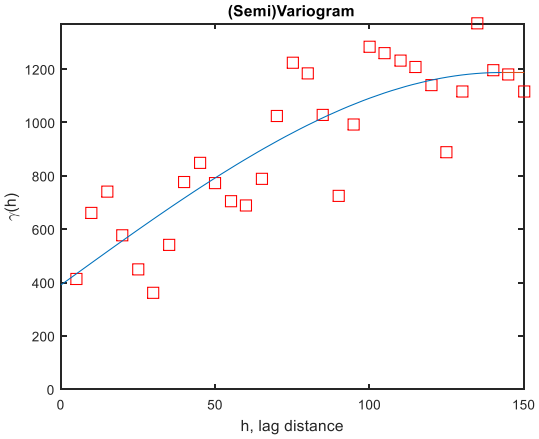


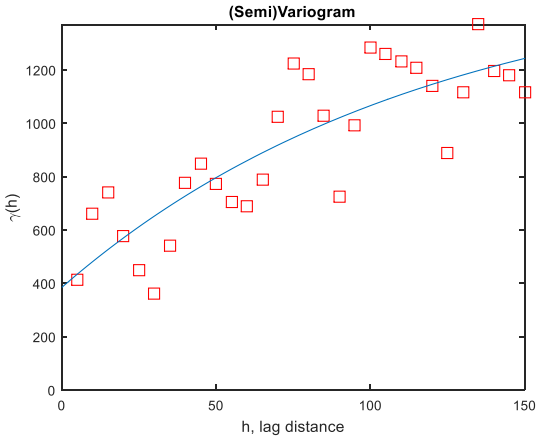
15m Shear wave Velocity with All Data

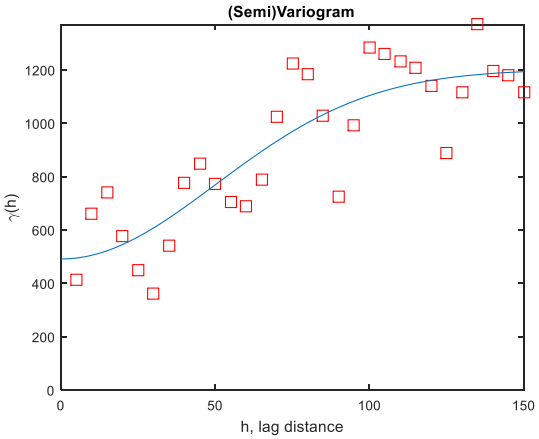




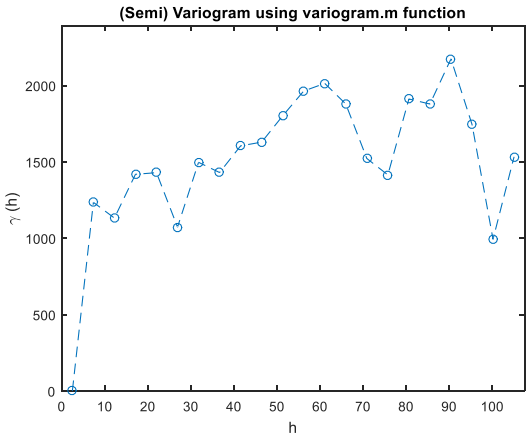


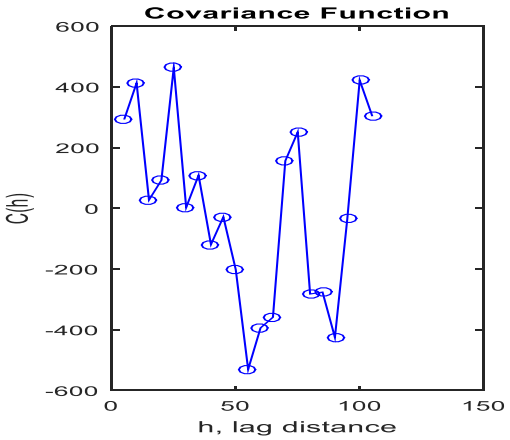


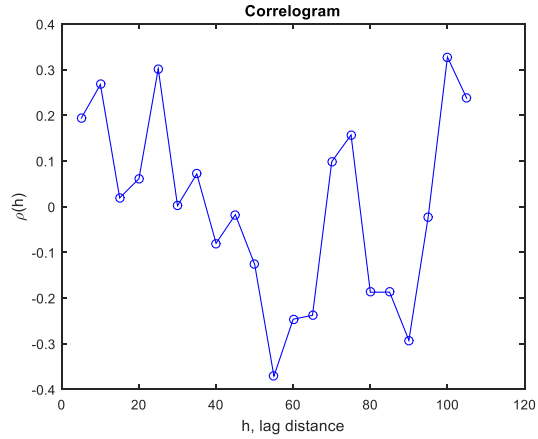


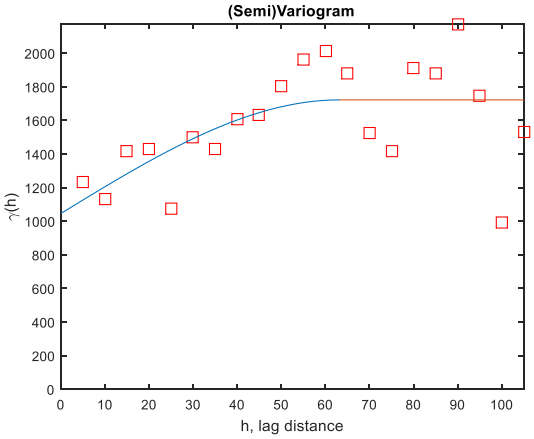


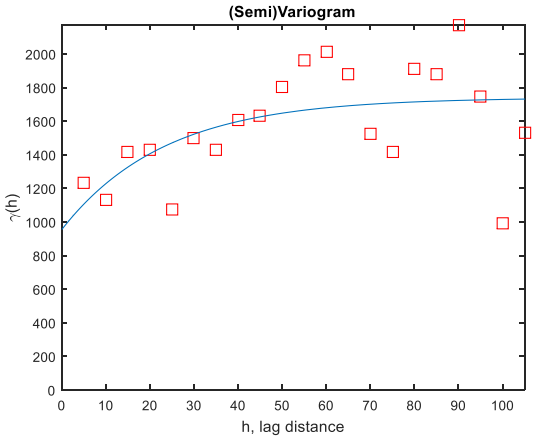
5m Rayleigh wave Velocity without Data after Hill

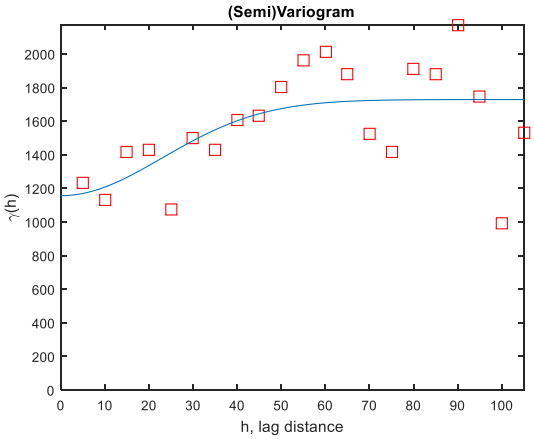




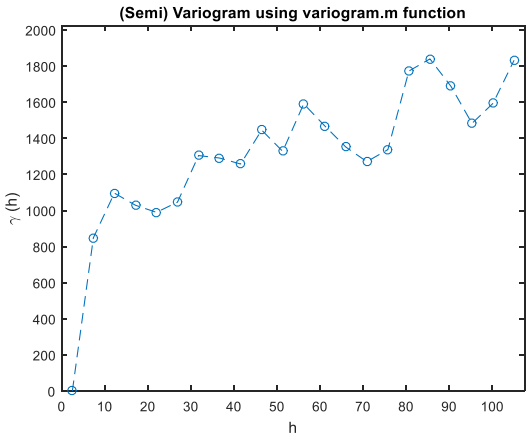


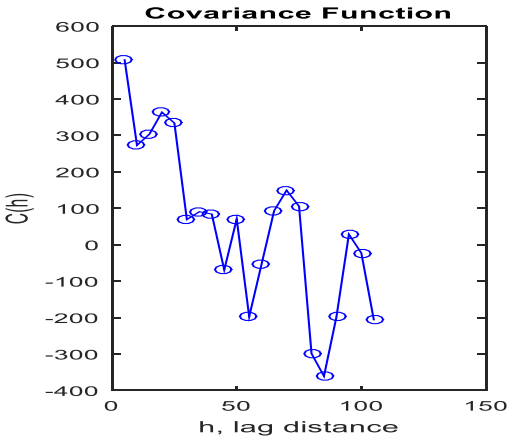


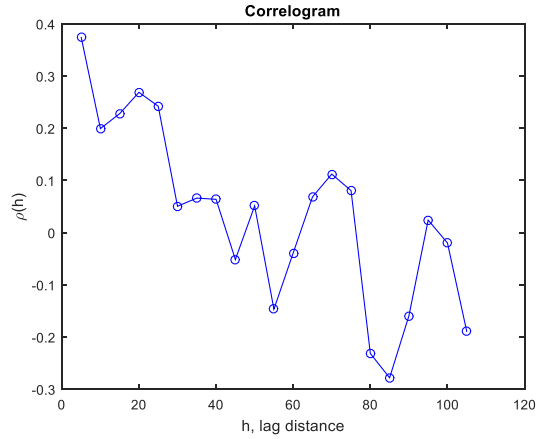


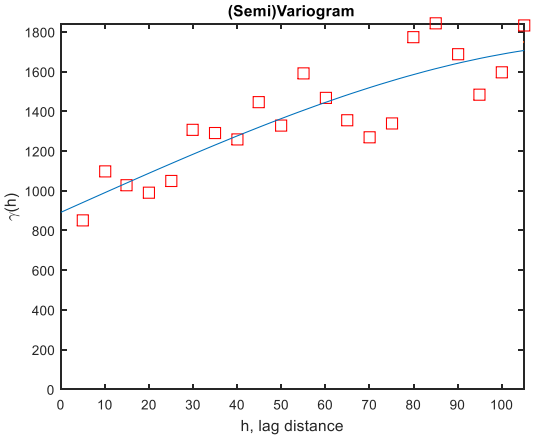


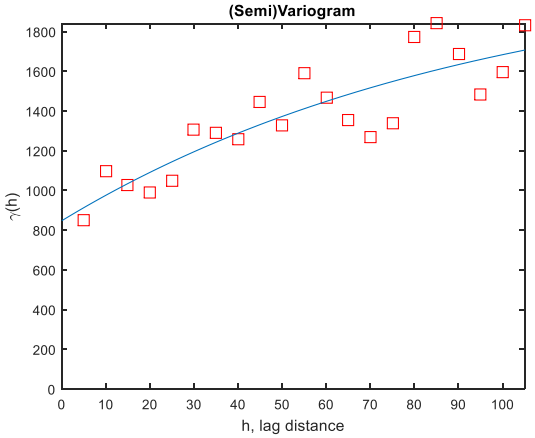
10m Rayleigh wave Velocity without Data after Hill

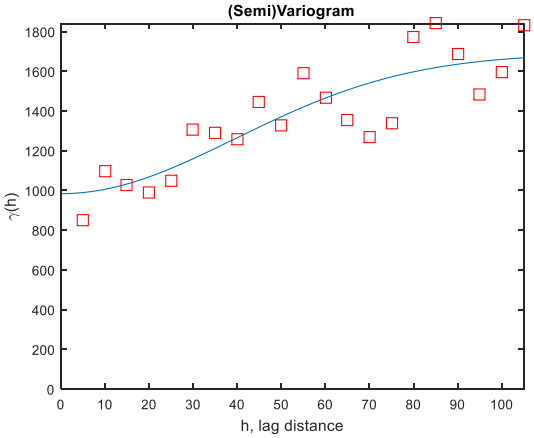




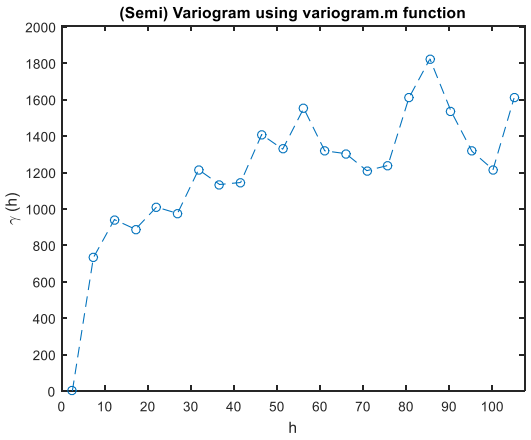


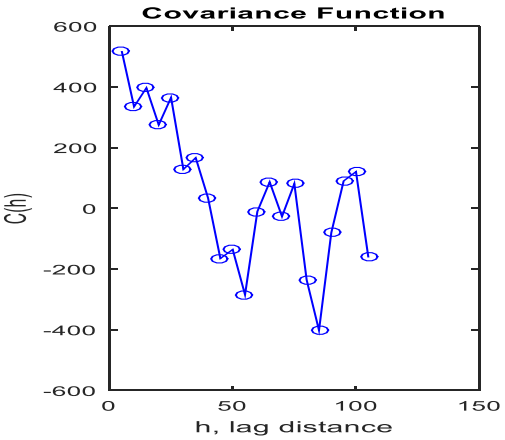


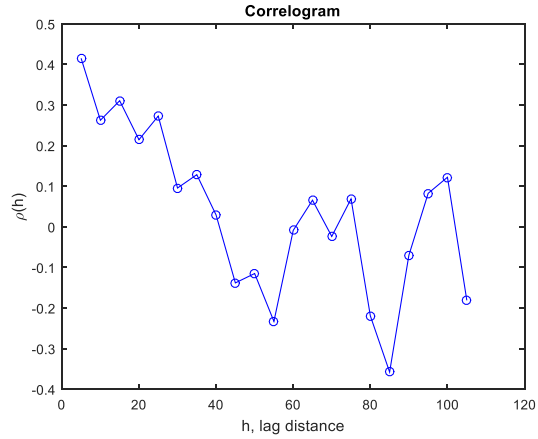


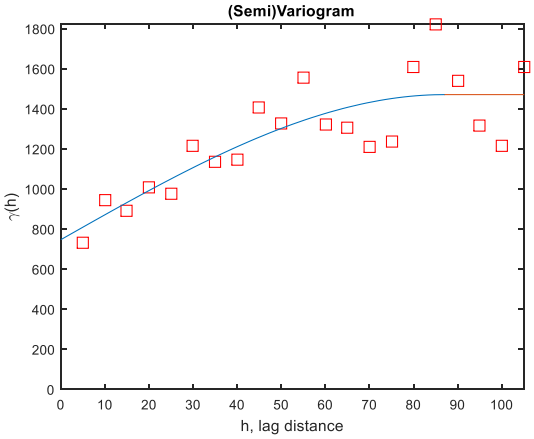


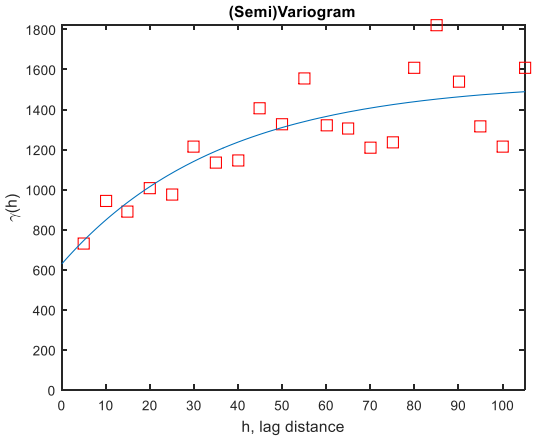
15m Rayleigh wave Velocity without Data after Hill

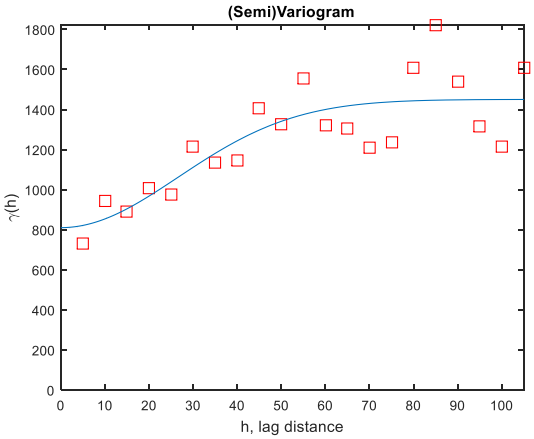




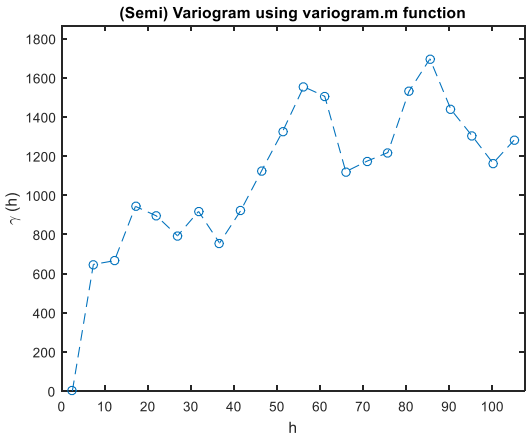


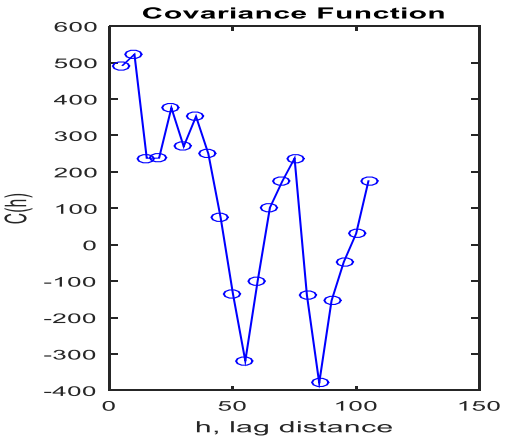


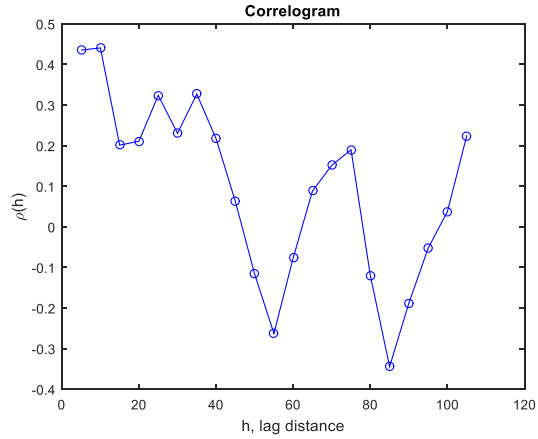


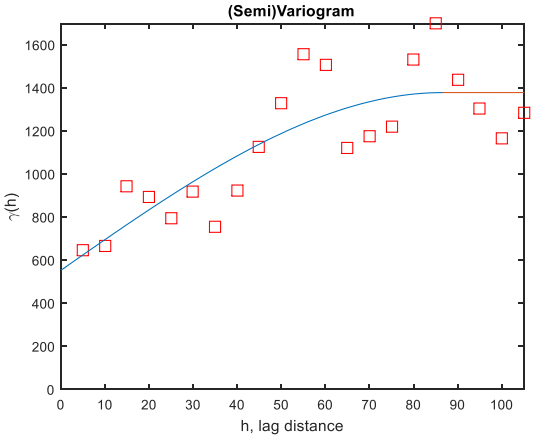


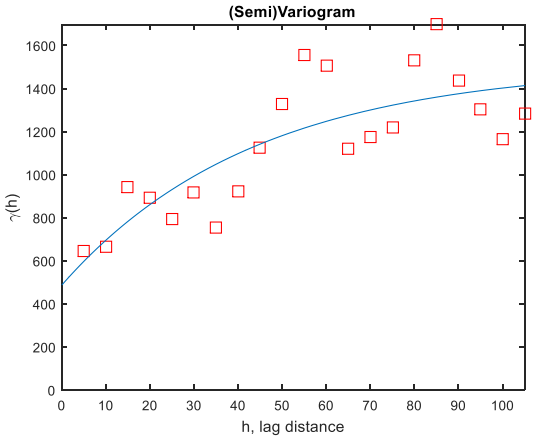
5m Shear wave Velocity without Data after Hill

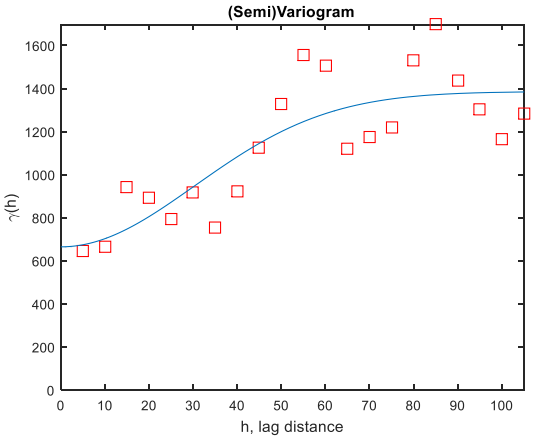




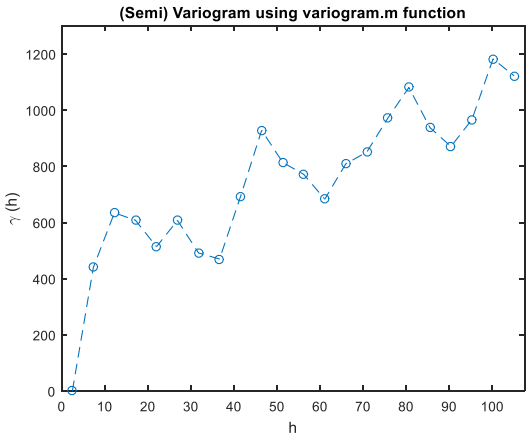


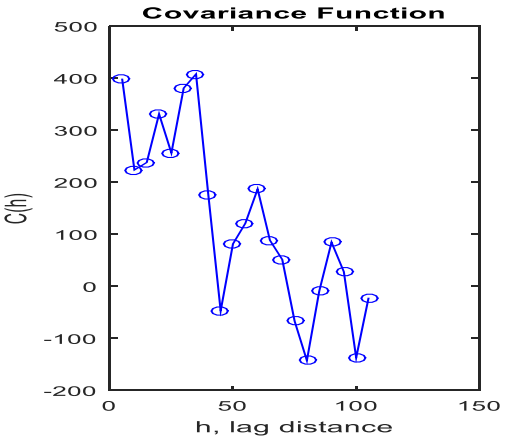


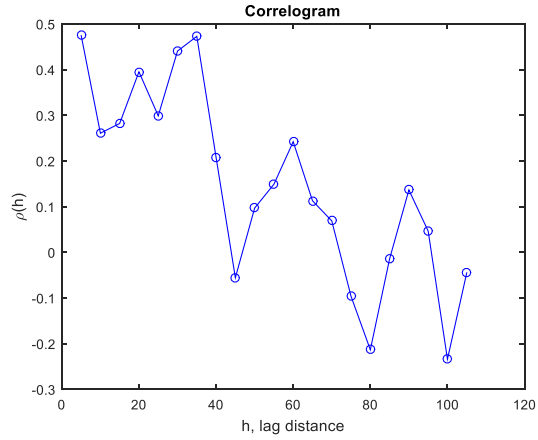


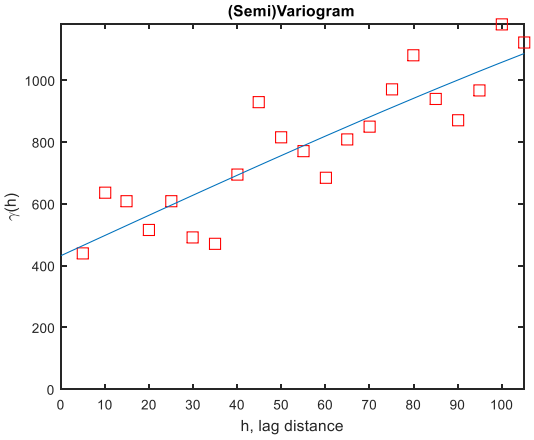


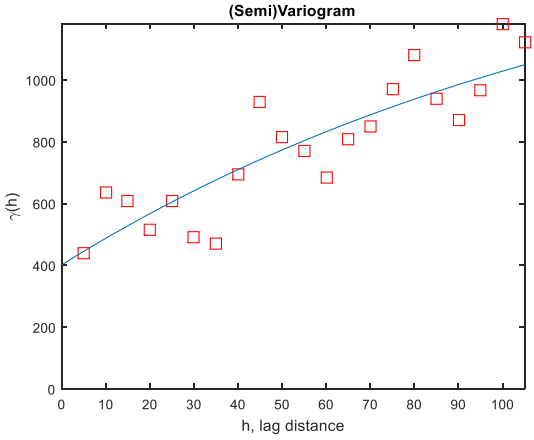
10m Shear wave Velocity without Data after Hill

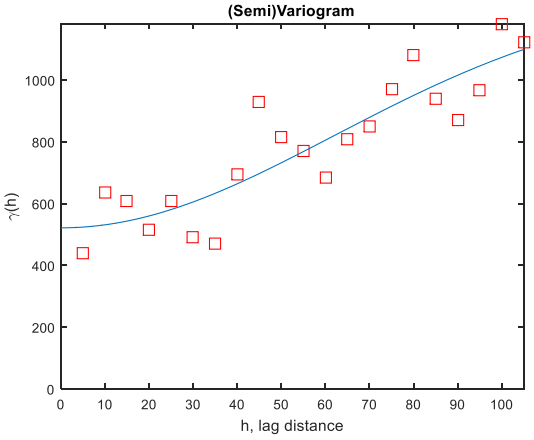




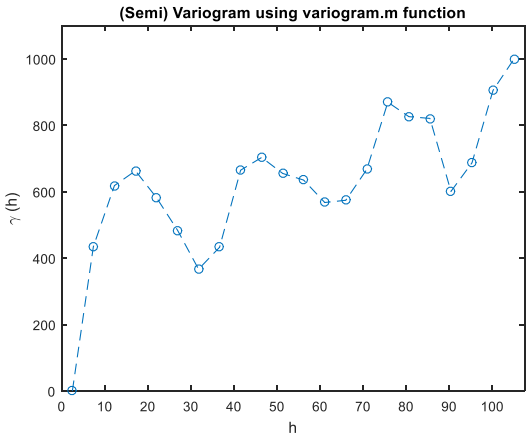


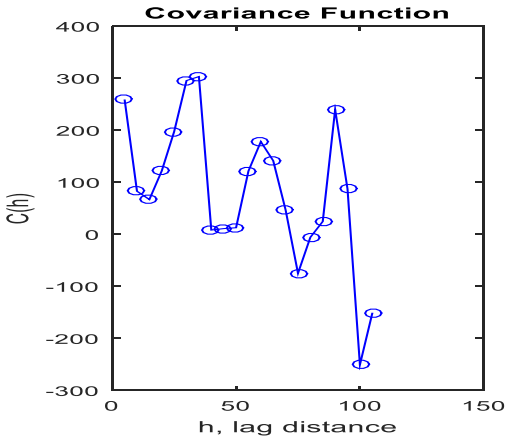


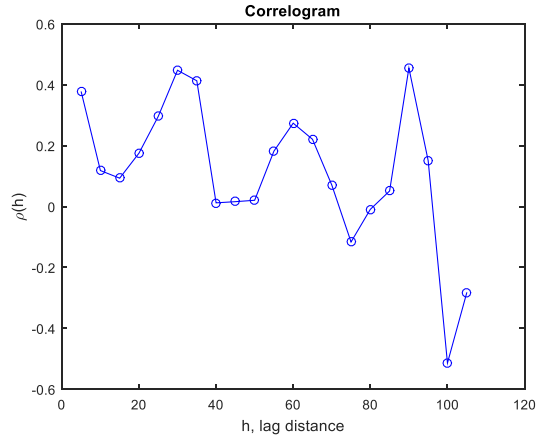


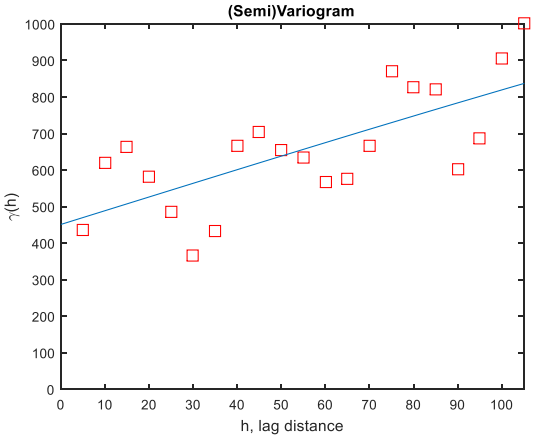


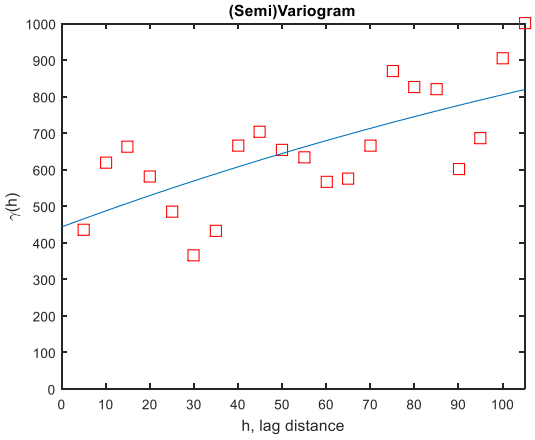
15m Shear wave Velocity without Data after Hill

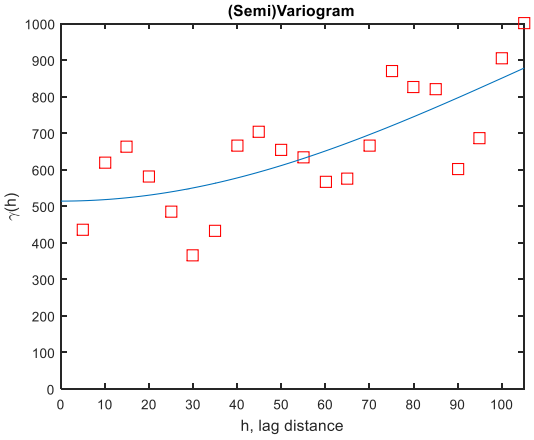




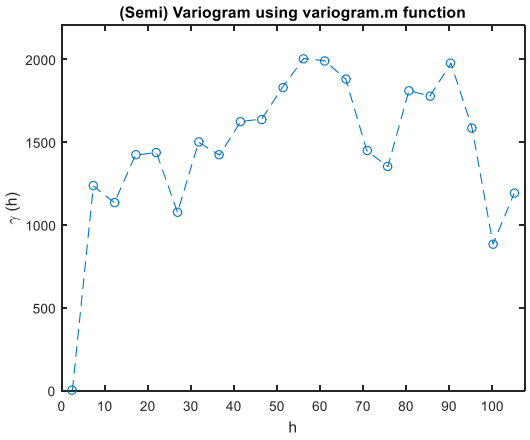


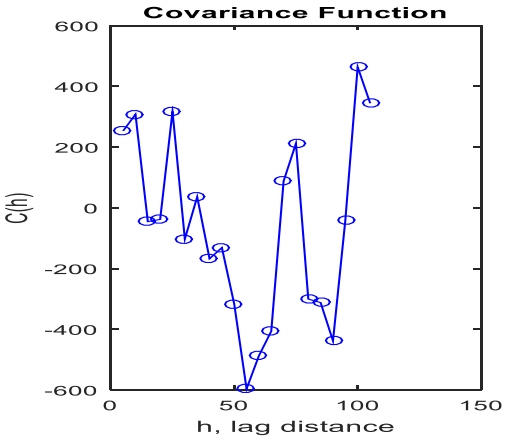


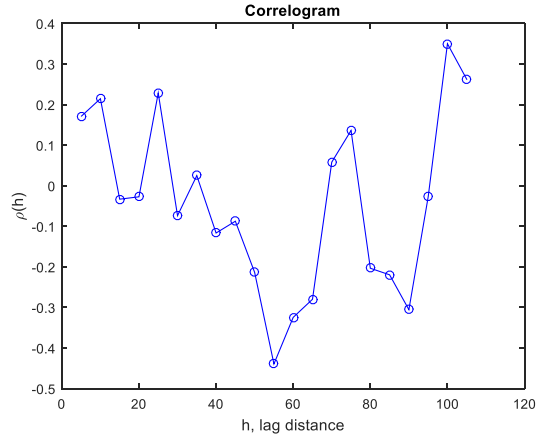


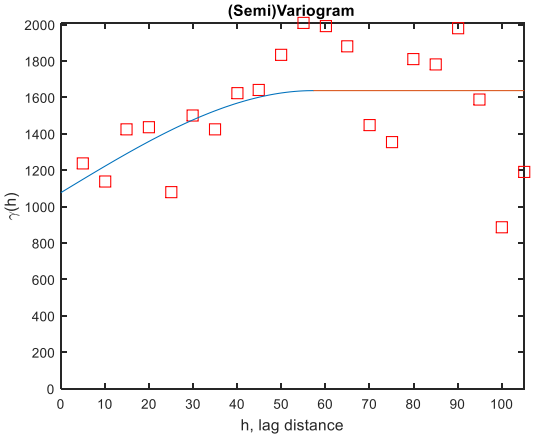


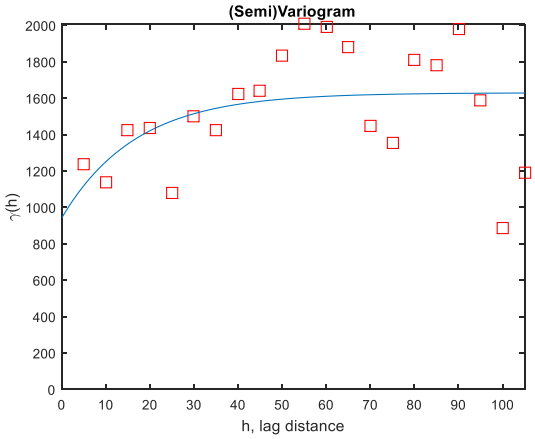
5m Rayleigh wave Velocity with Trend Removed as Linear Function without Data after Hill

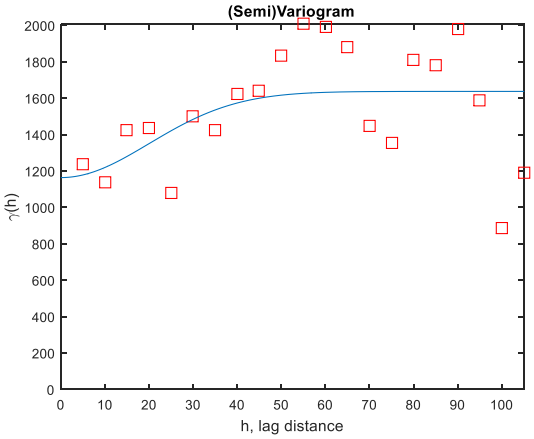




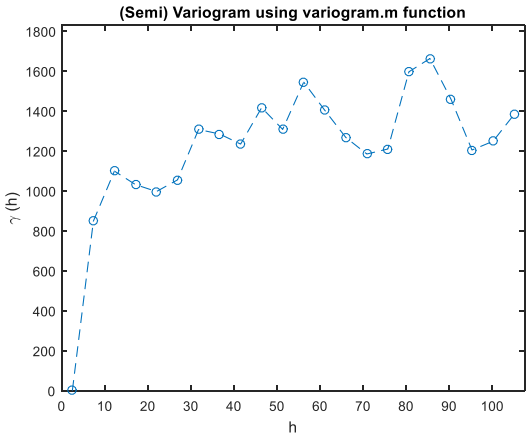


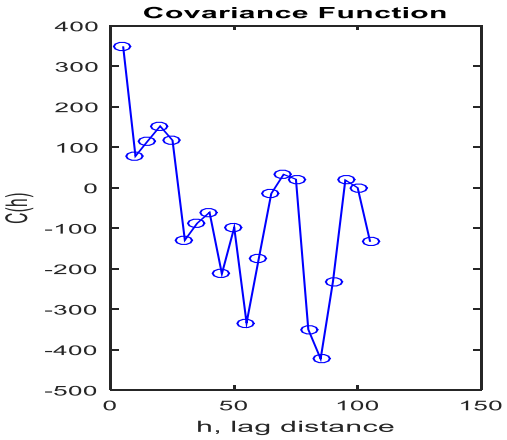


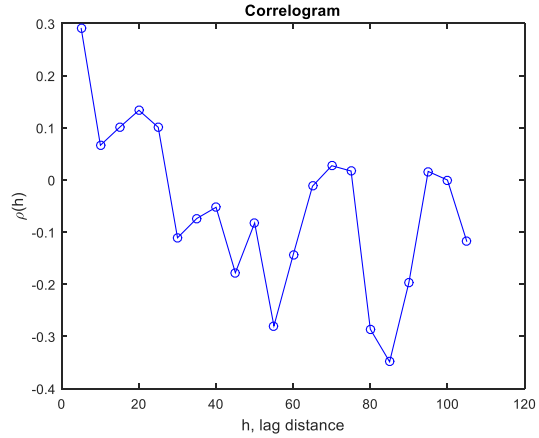


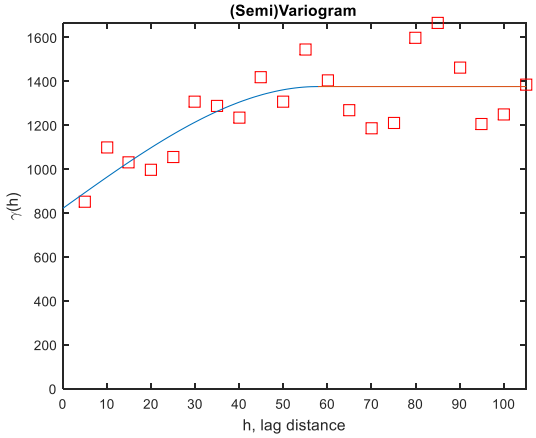


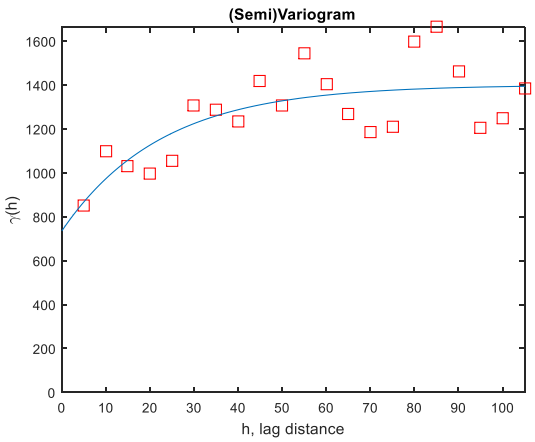
10m Rayleigh wave Velocity with Trend Removed as Linear Function without Data after Hill

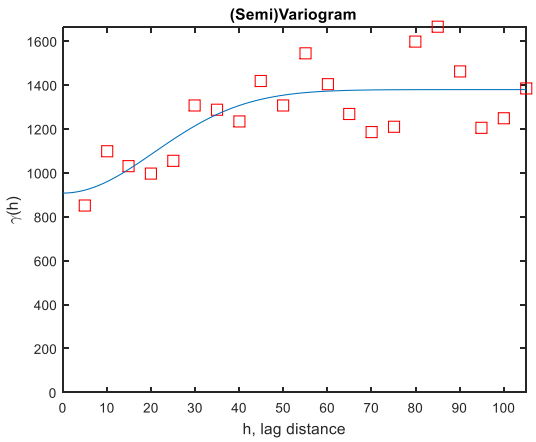




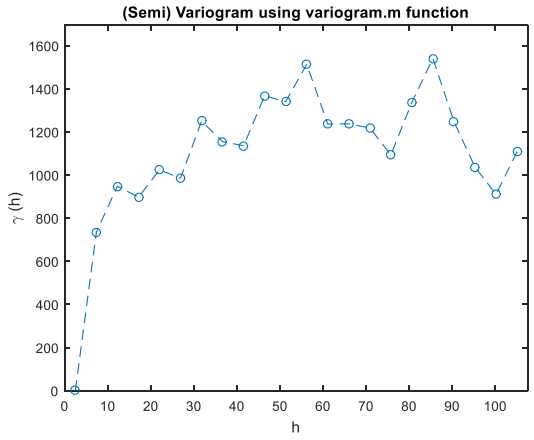


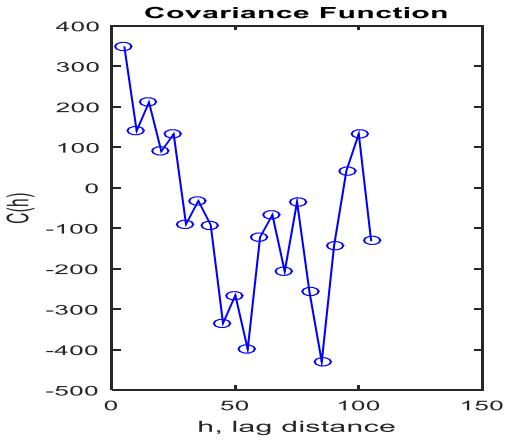


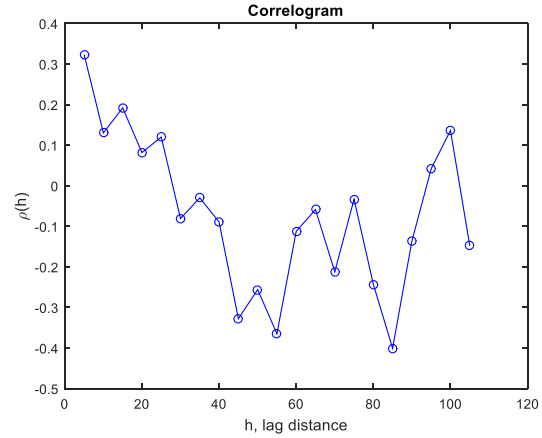


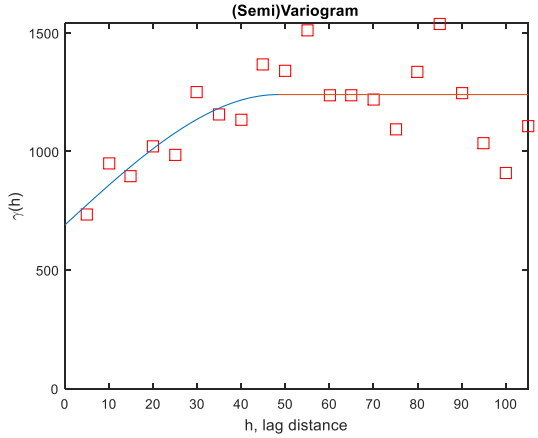


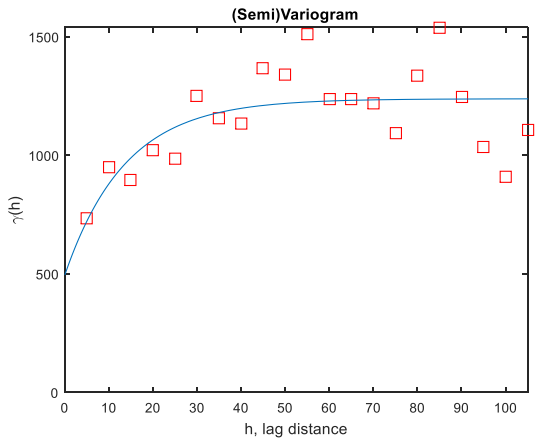
15m Rayleigh wave Velocity with Trend Removed as Linear Function without Data after Hill

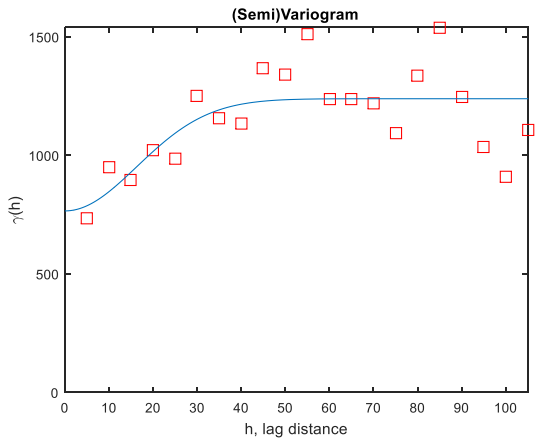




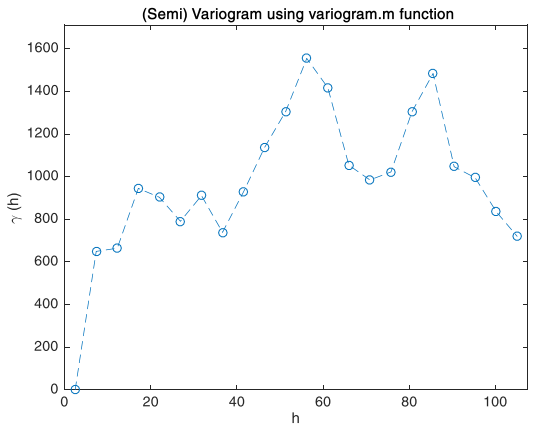


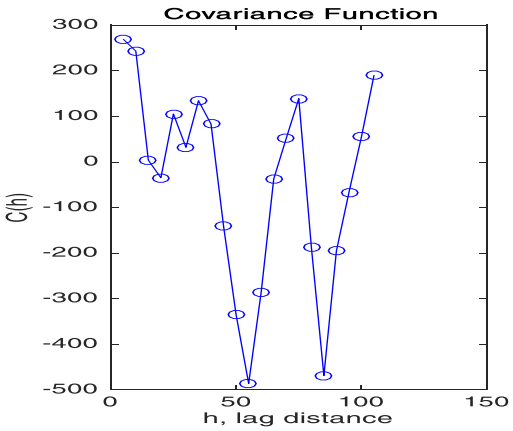


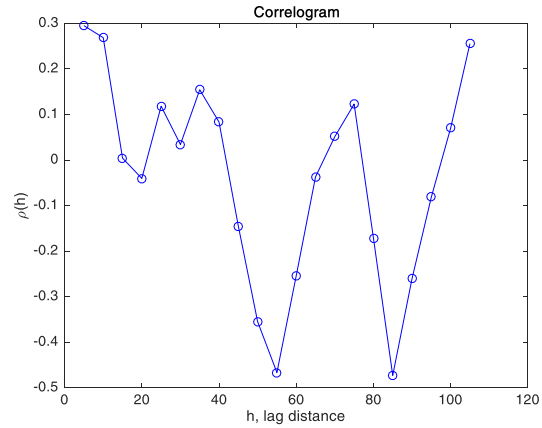


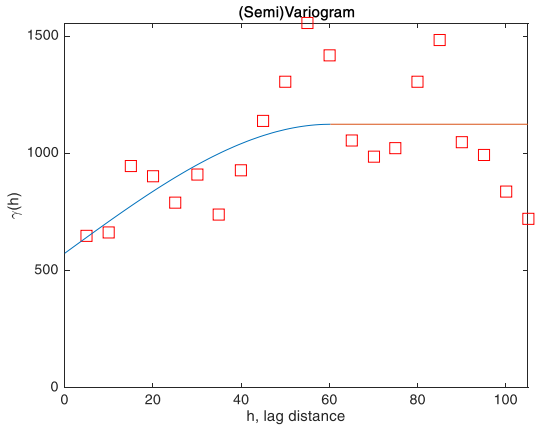


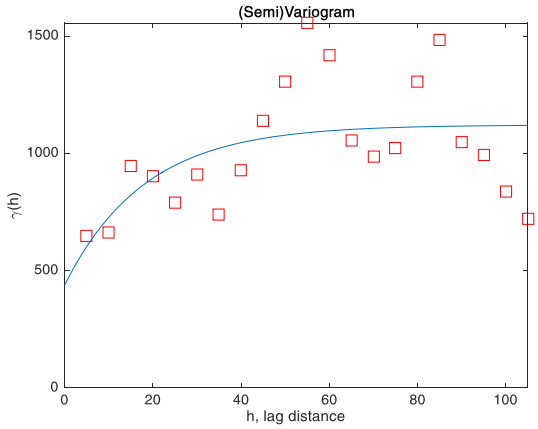
5m Shear wave Velocity with Trend Removed as Linear Function without Data after Hill

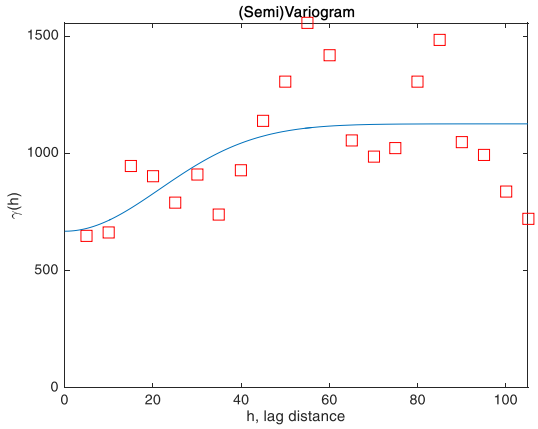




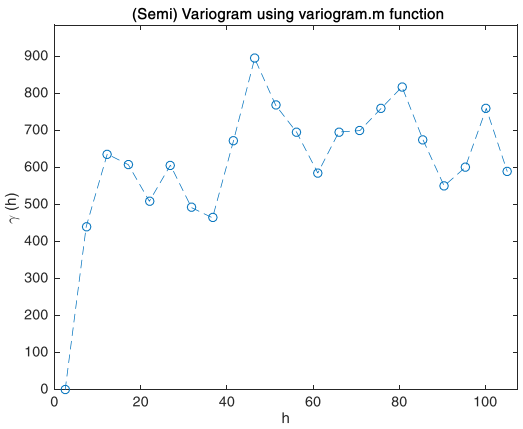


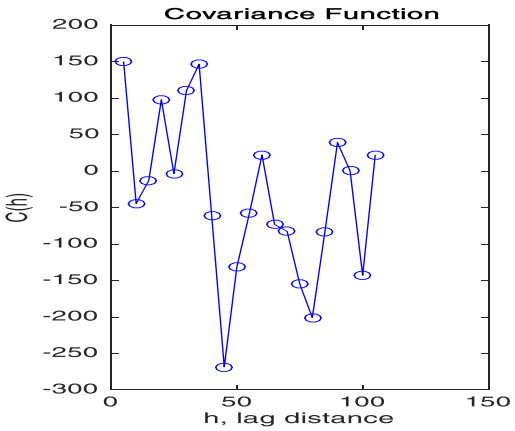


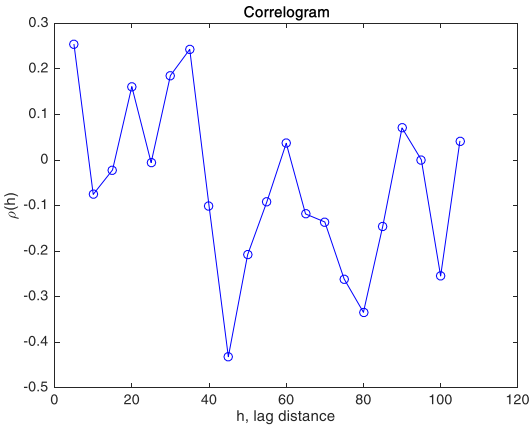


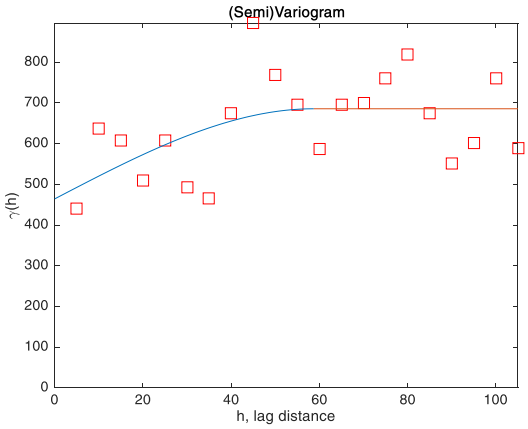


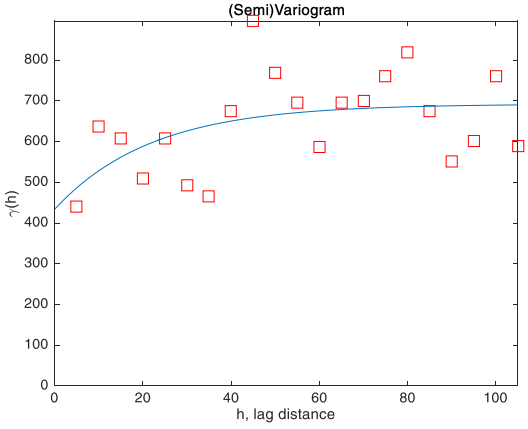
10m Shear wave Velocity with Trend Removed as Linear Function without Data after Hill

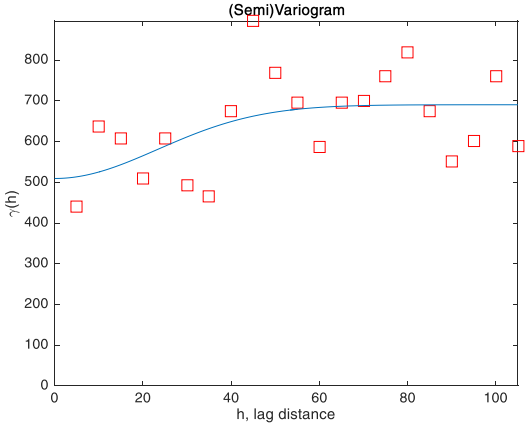




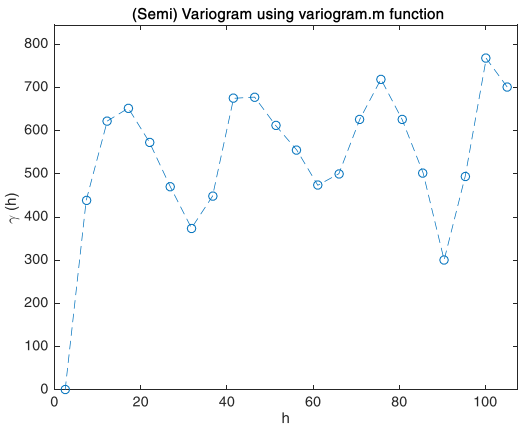


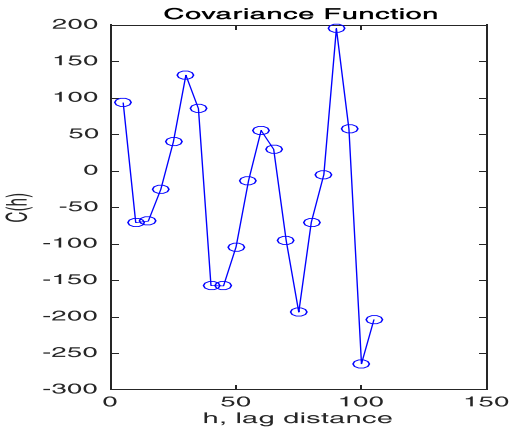


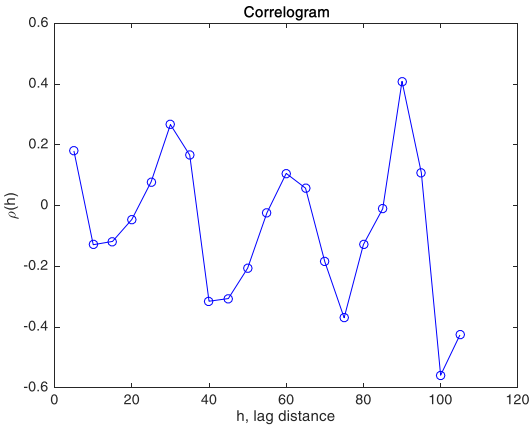


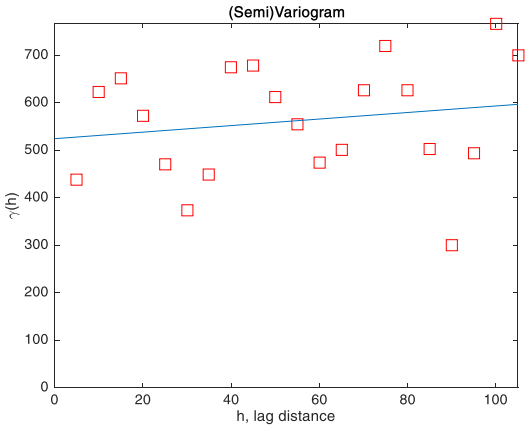


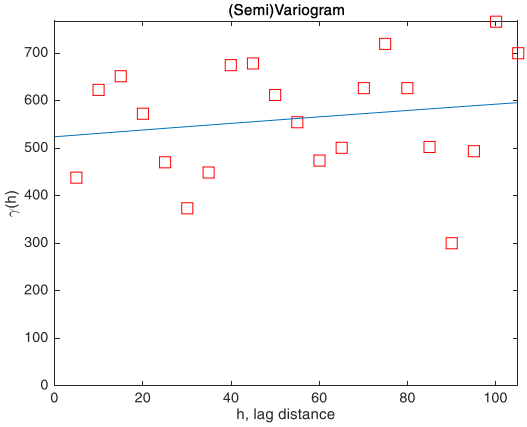
15m Shear wave Velocity with Trend Removed as Linear Function without Data after Hill

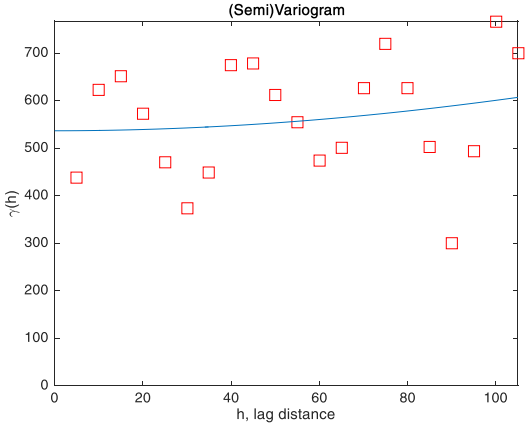




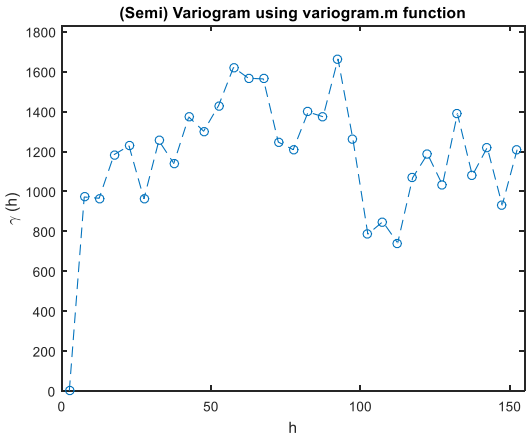


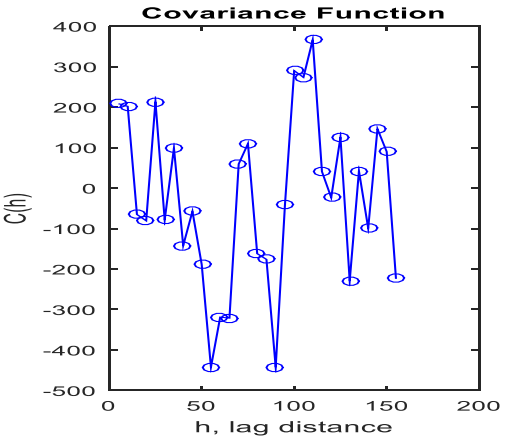


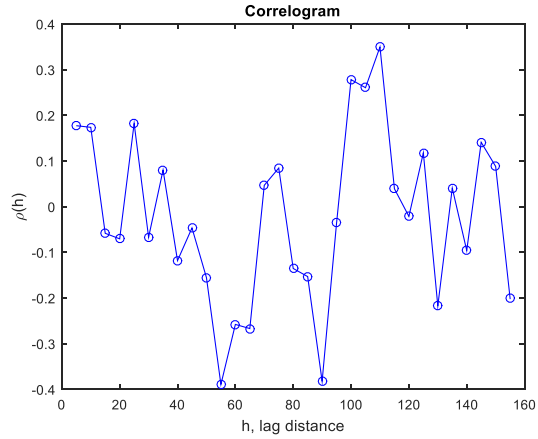


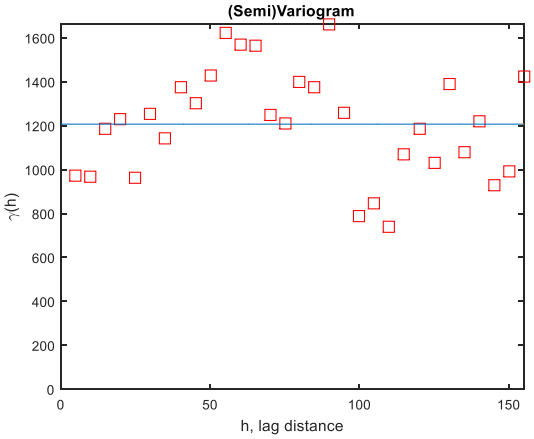


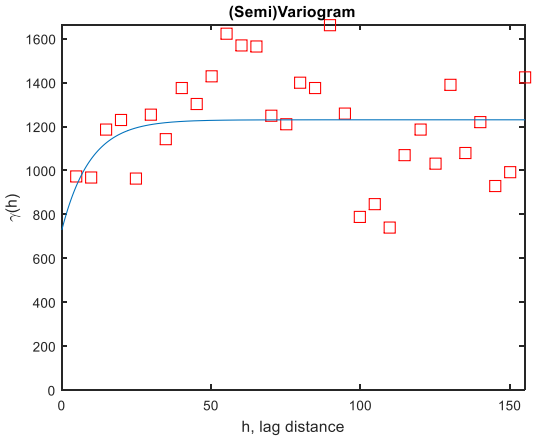
5m Rayleigh wave Velocity with Trend Removed as Linear Piecewise Function

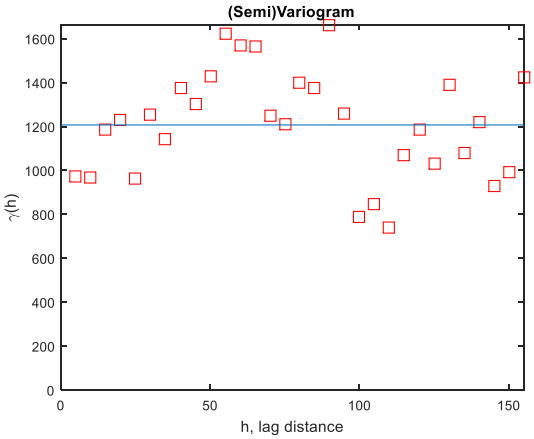




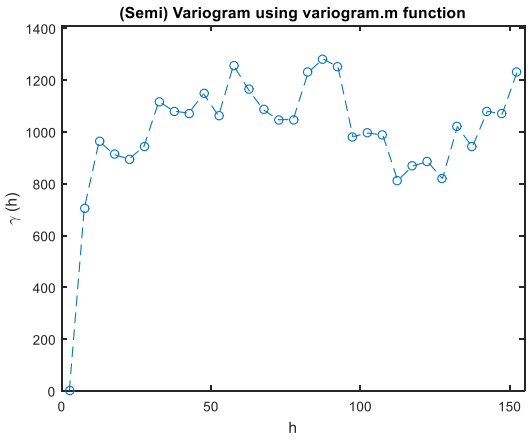


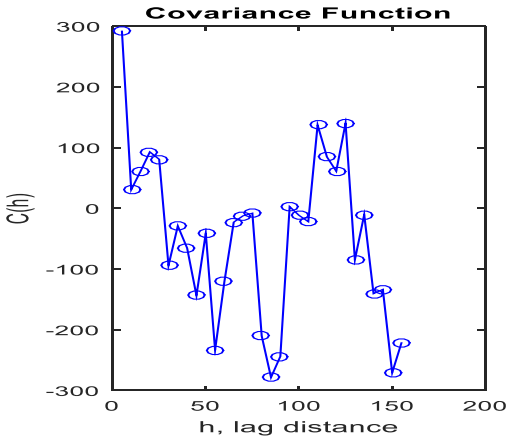


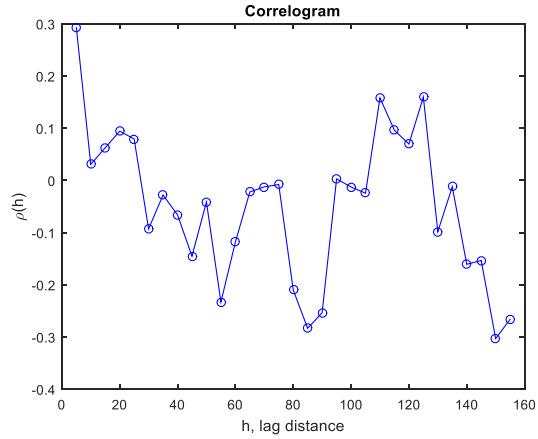


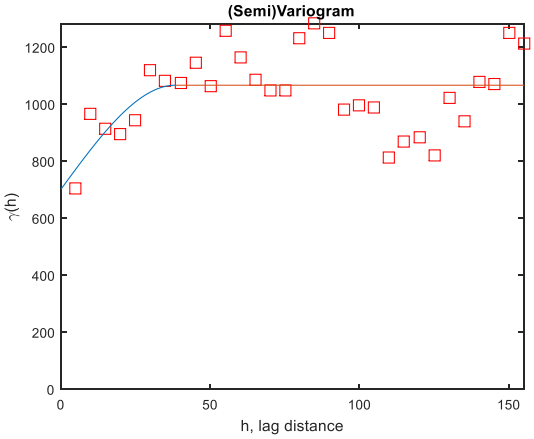


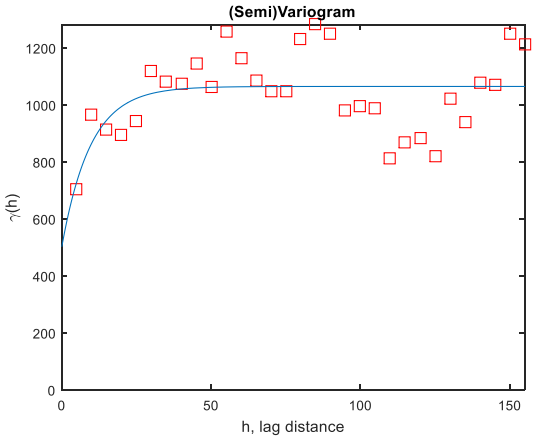
10m Rayleigh wave Velocity with Trend Removed as Linear Piecewise Function

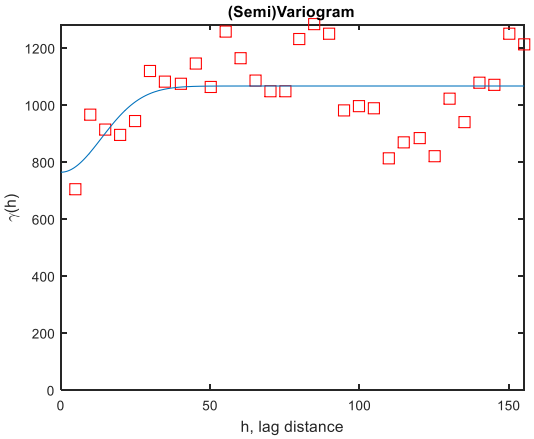




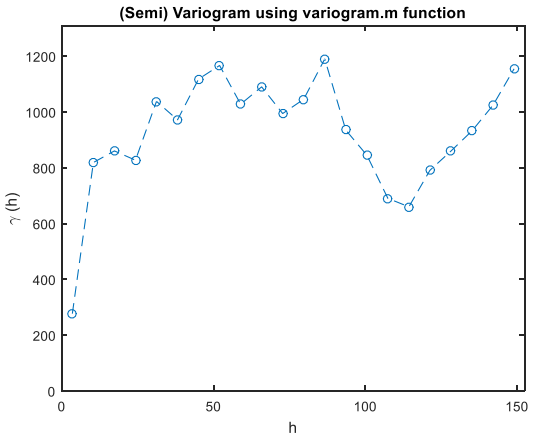


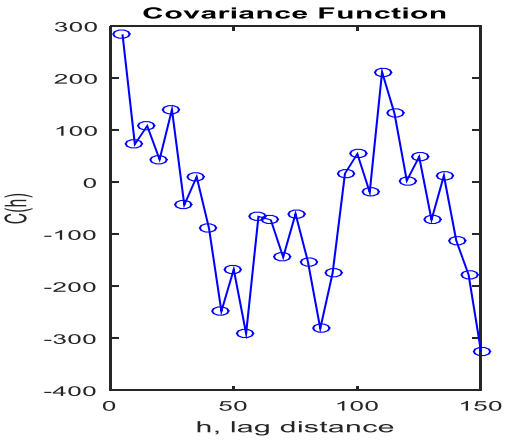


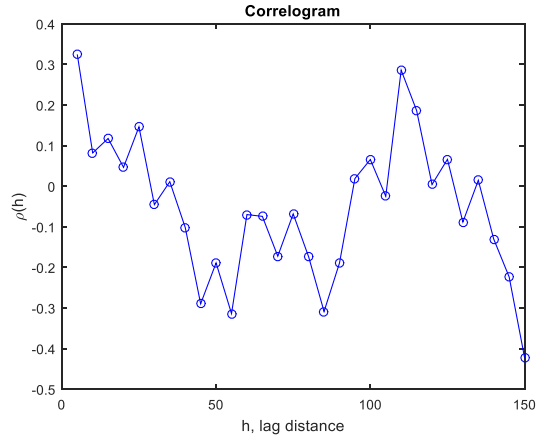


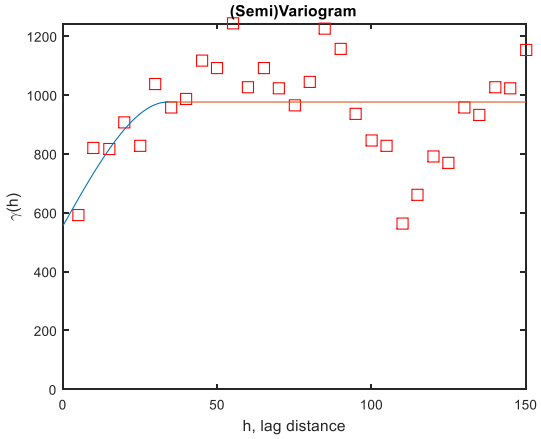


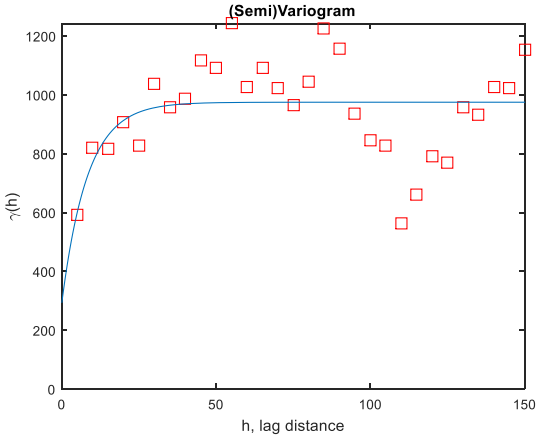
15m Rayleigh wave Velocity with Trend Removed as Linear Piecewise Function

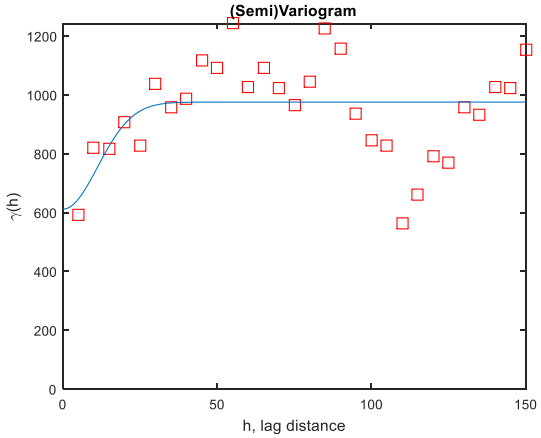




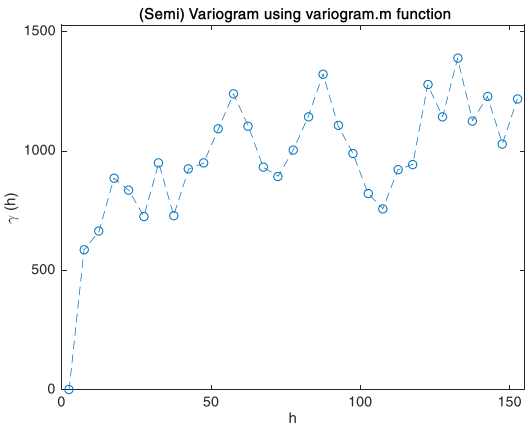


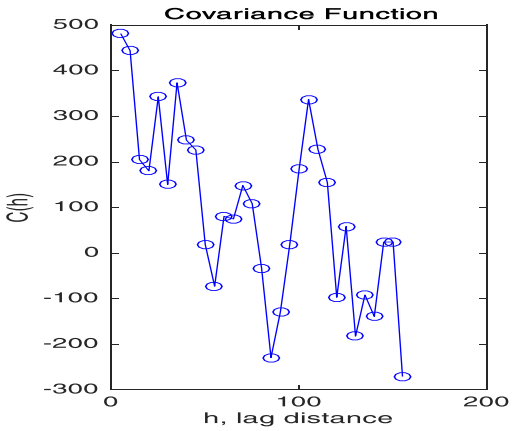


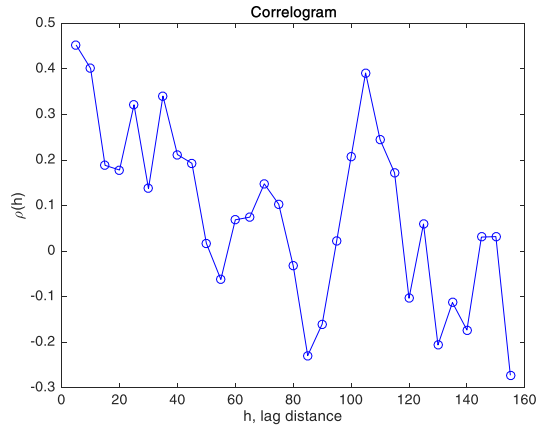


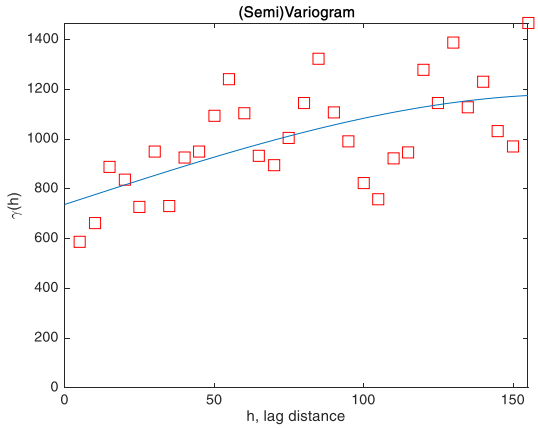


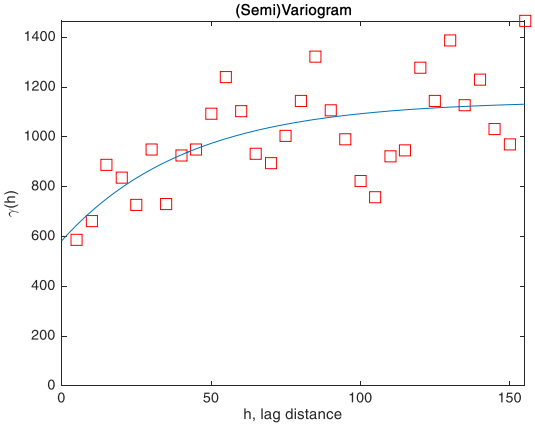
5m Shear wave Velocity with Trend Removed as Linear Piecewise Function

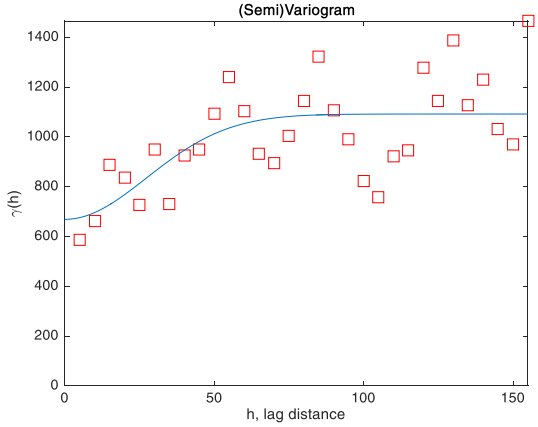




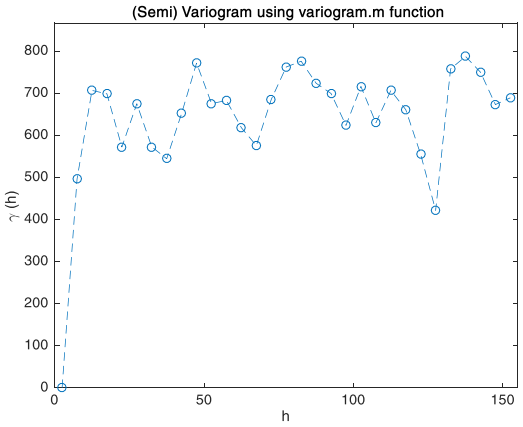


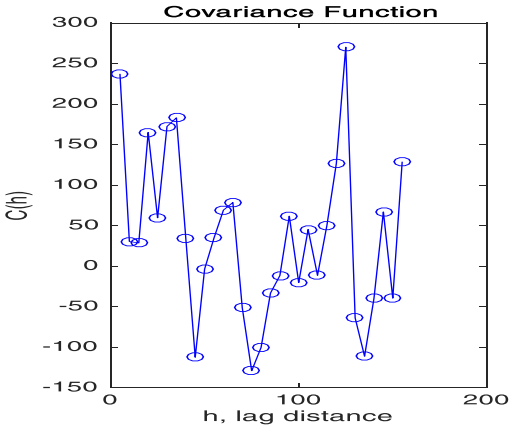


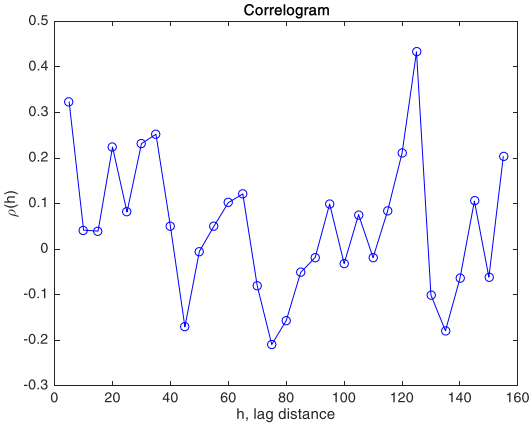


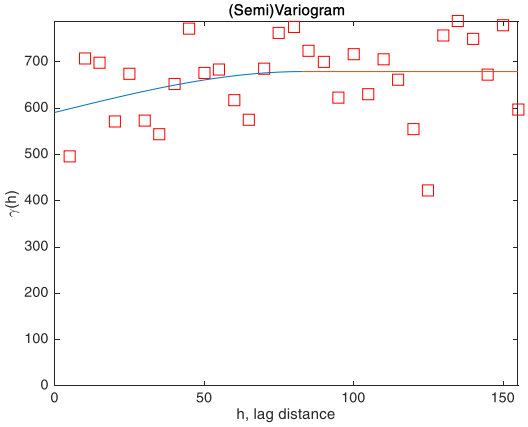


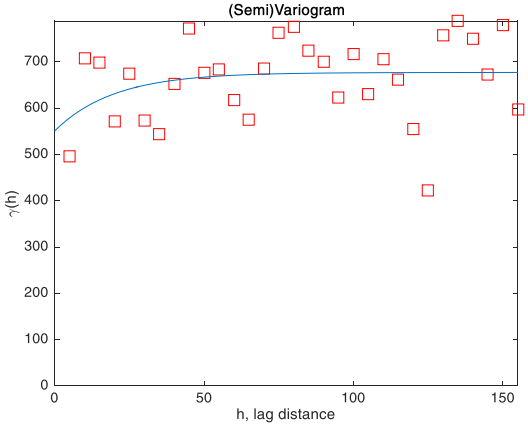
10m Shear wave Velocity with Trend Removed as Linear Piecewise Function

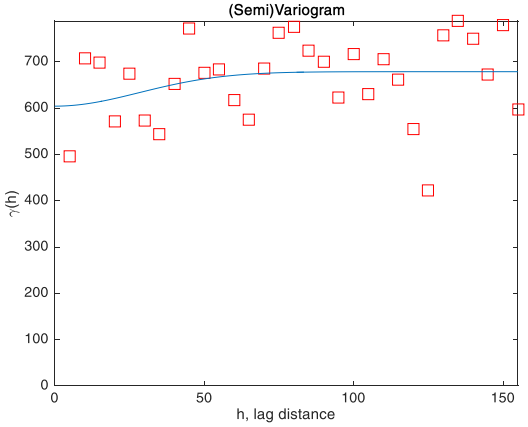




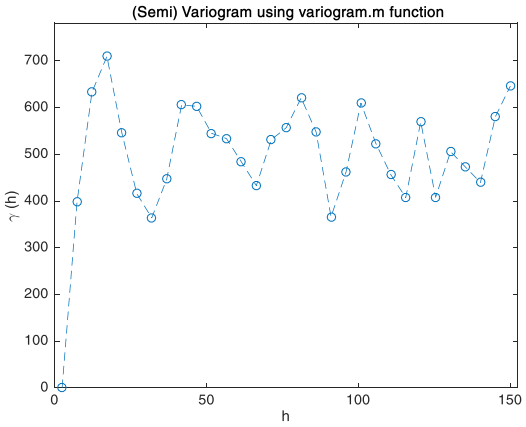


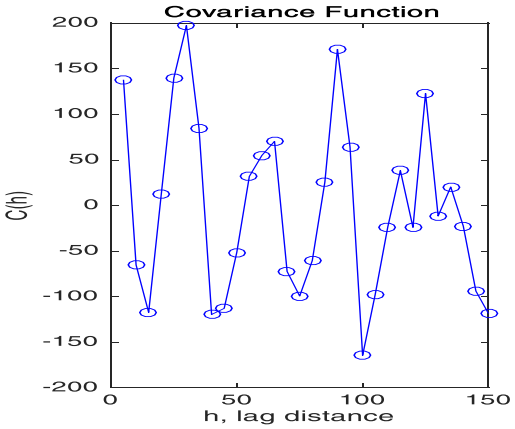


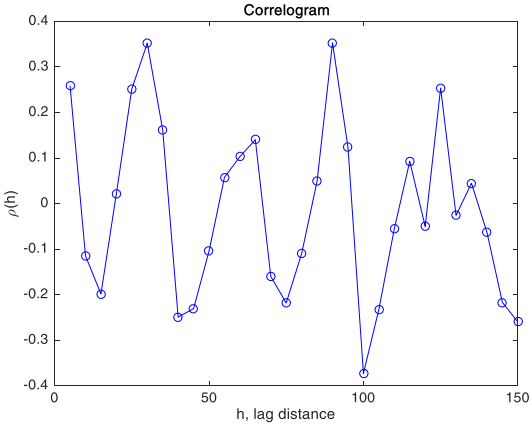


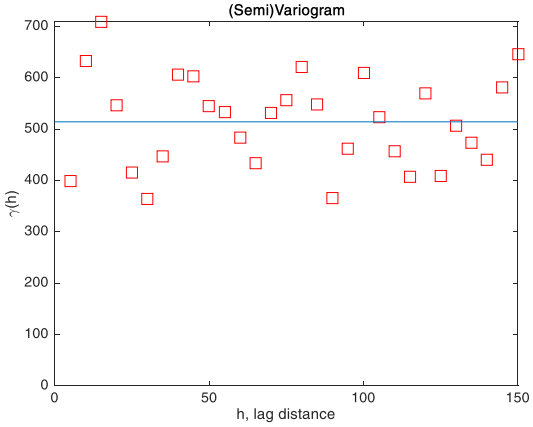


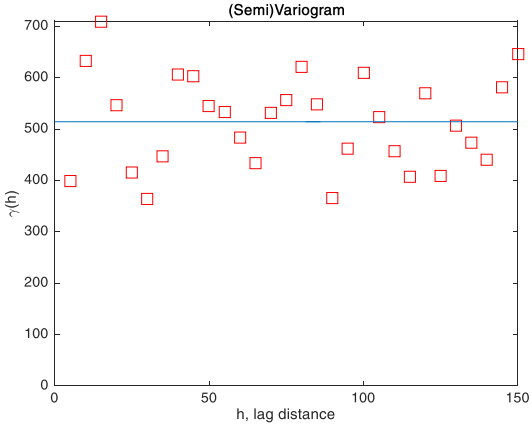
15m Shear wave Velocity with Trend Removed as Linear Piecewise Function

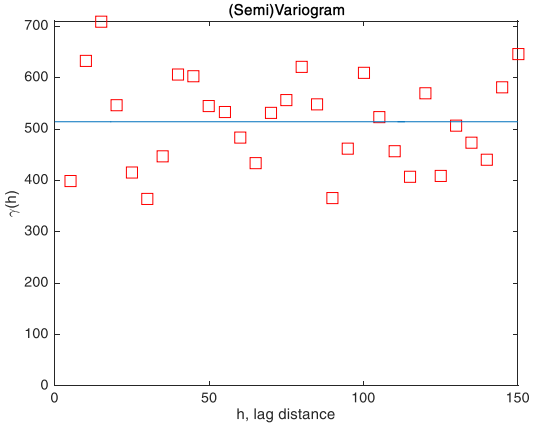




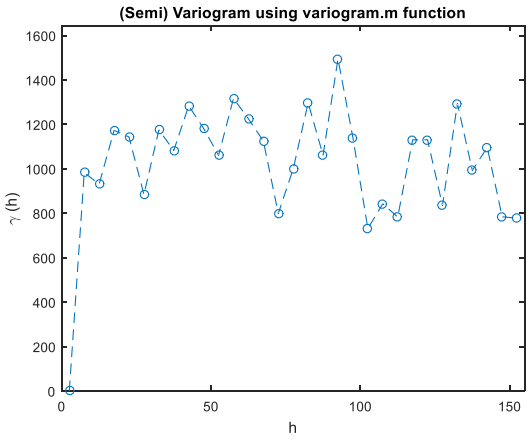


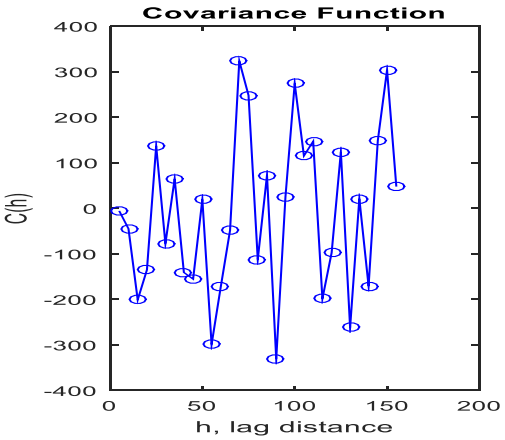


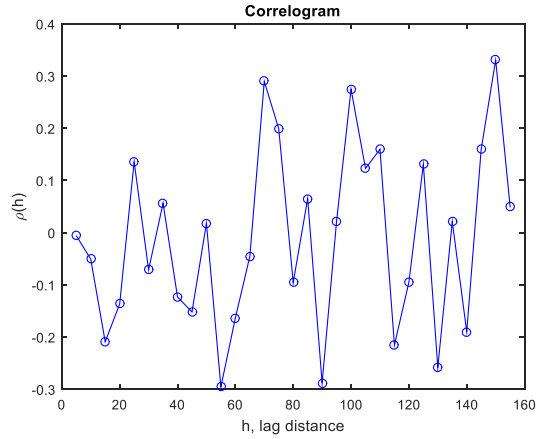


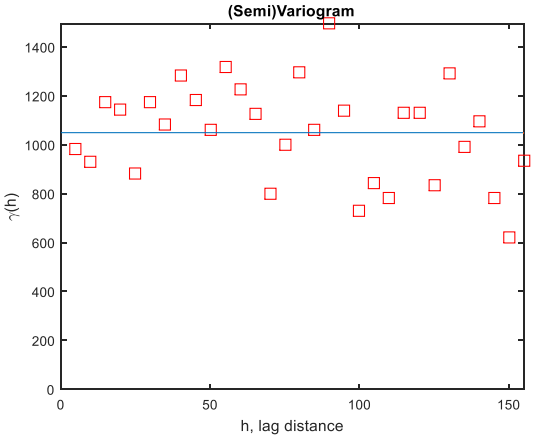


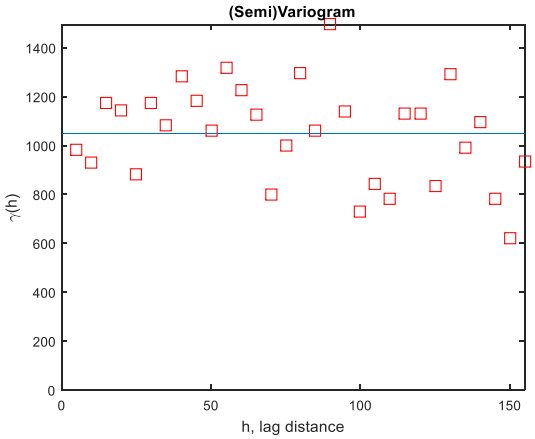
5m Rayleigh wave Velocity with Trend Removed as 4th-Order Polynomial Function

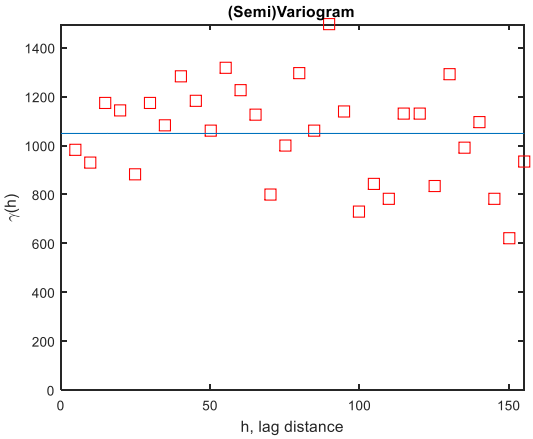




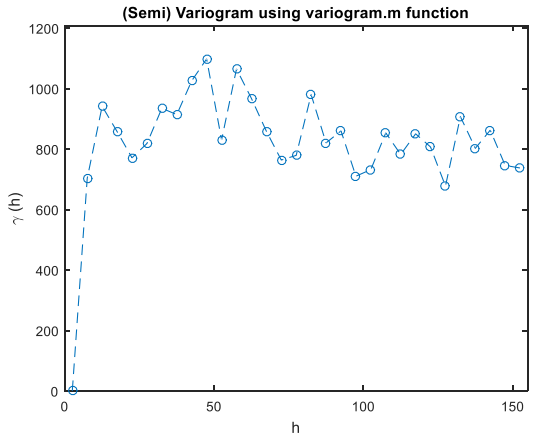


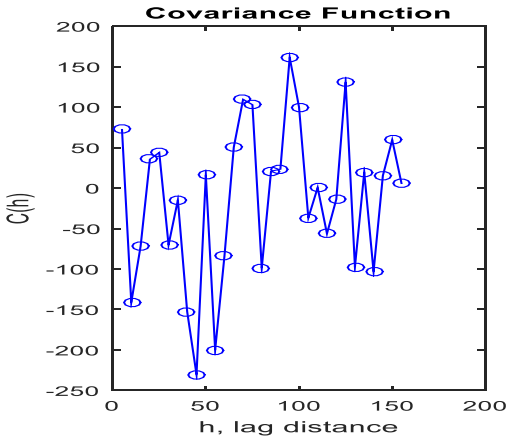


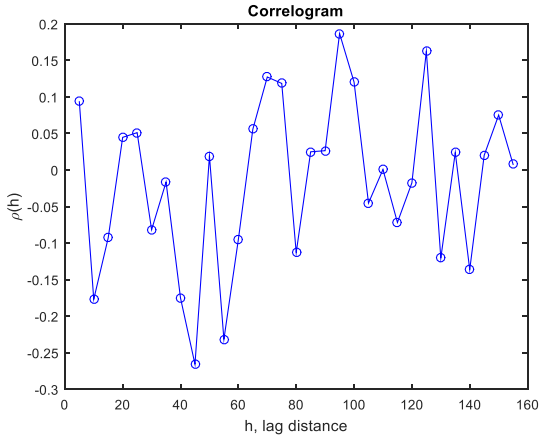


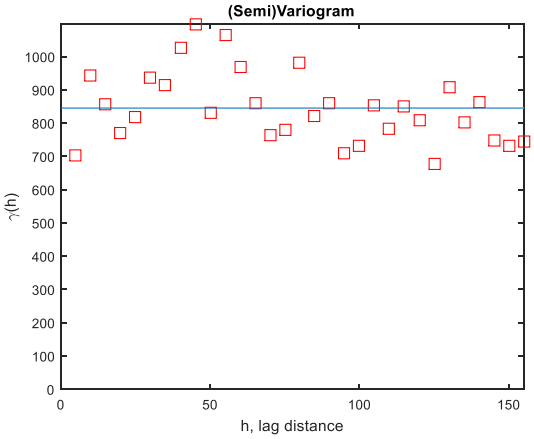


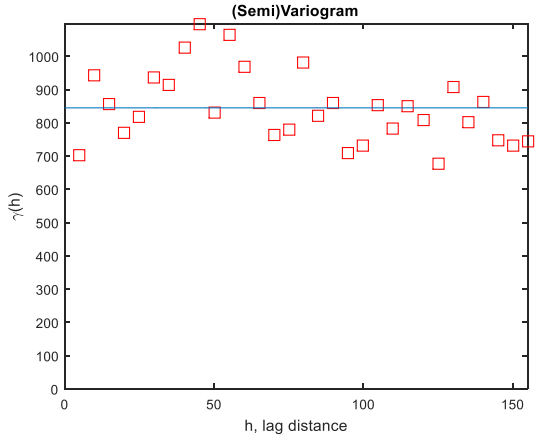
10m Rayleigh wave Velocity with Trend Removed as 4th-Order Polynomial Function

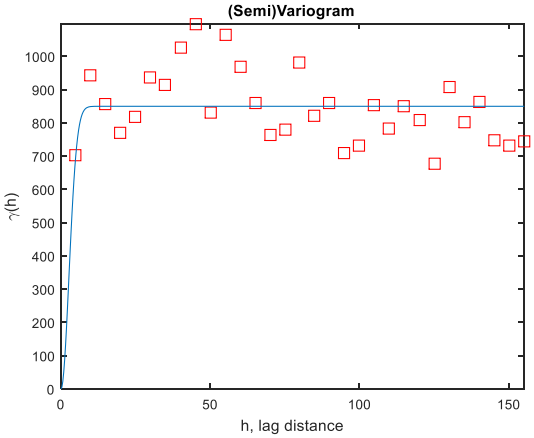




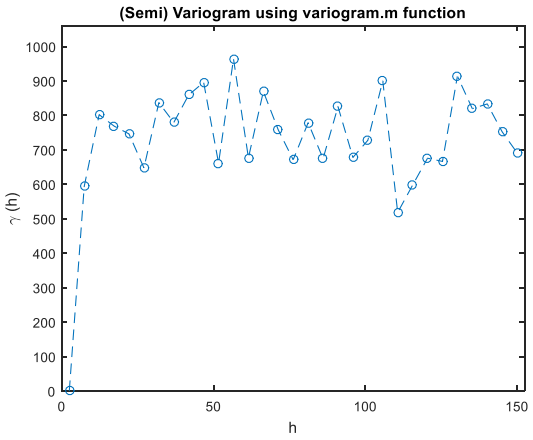


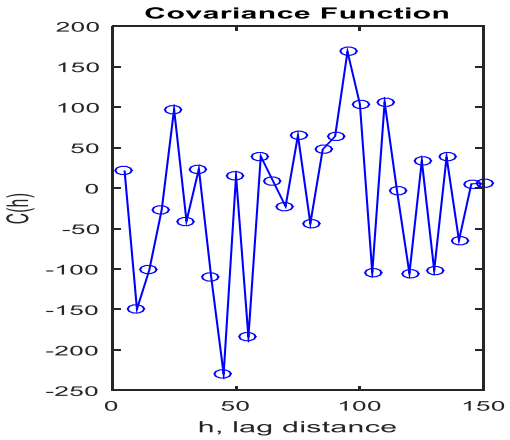


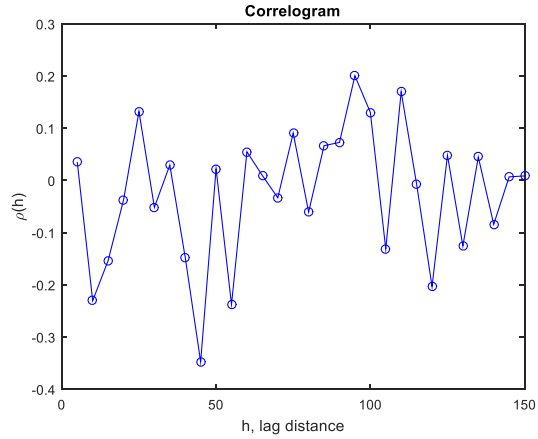


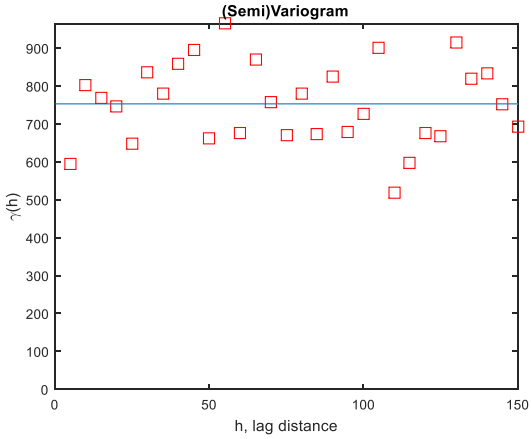


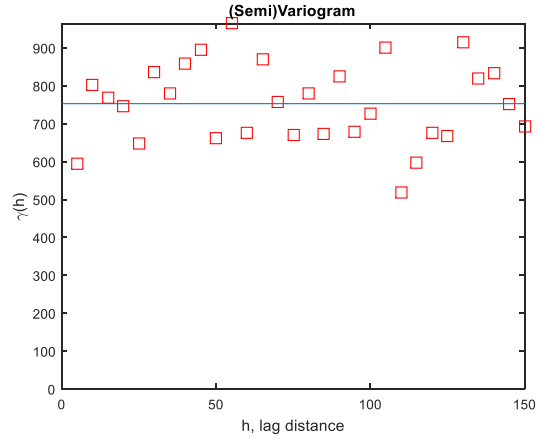
15m Rayleigh wave Velocity with Trend Removed as 4th-Order Polynomial Function

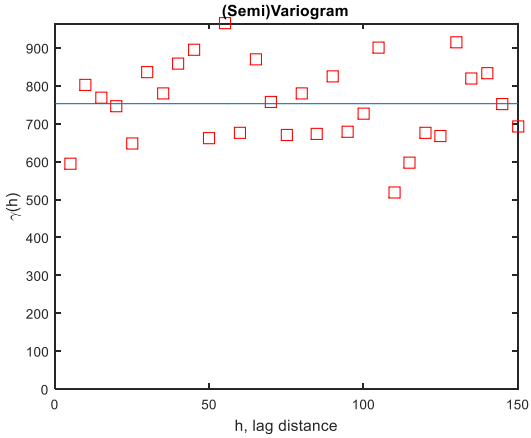




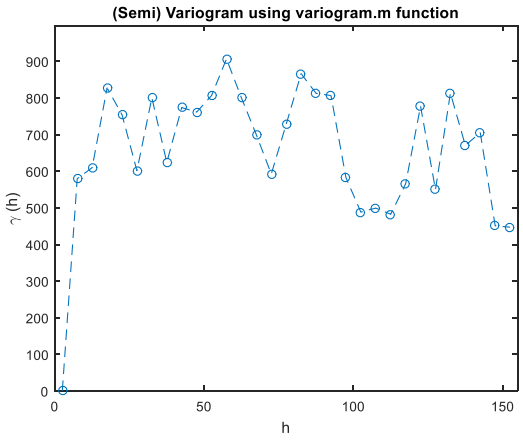


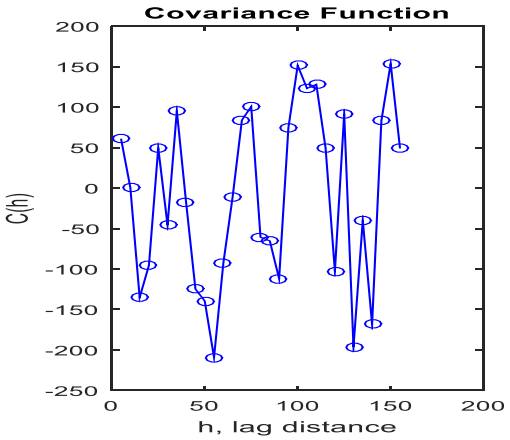


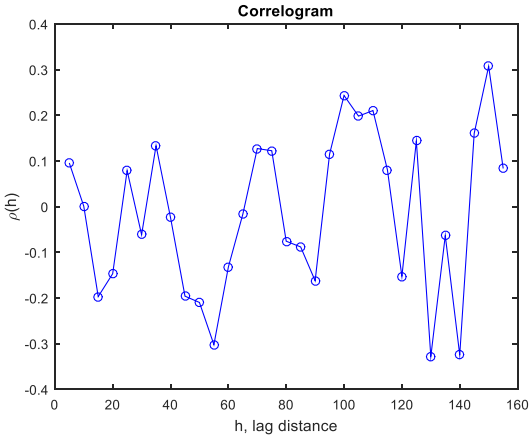


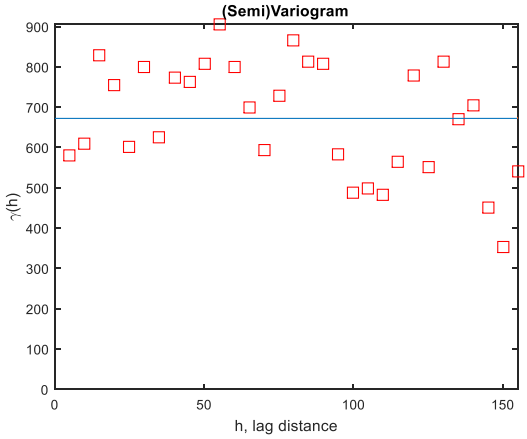


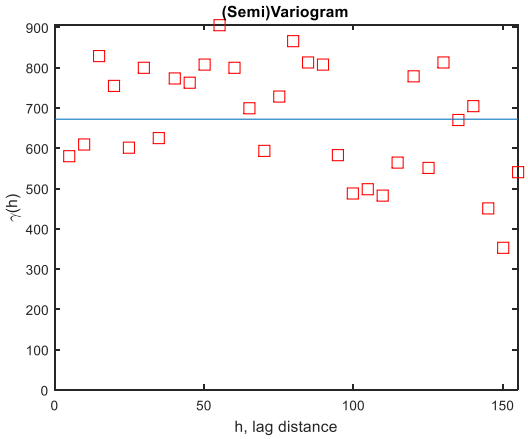
5m Shear wave Velocity with Trend Removed as 4th-Order Polynomial Function

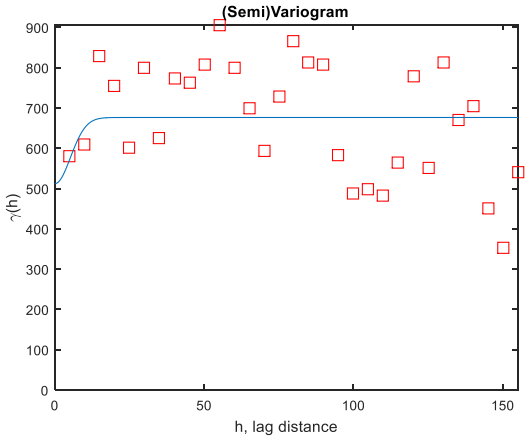




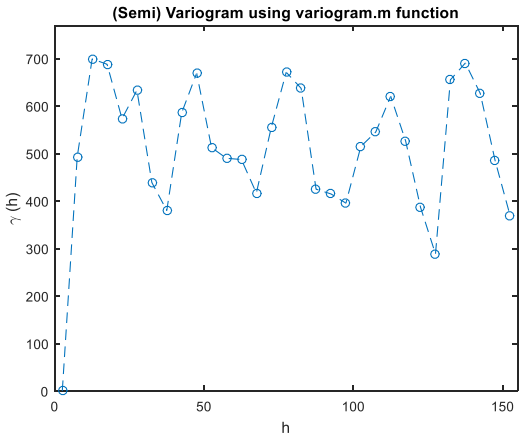


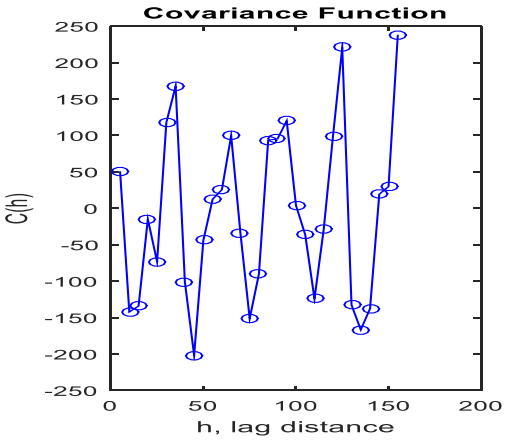


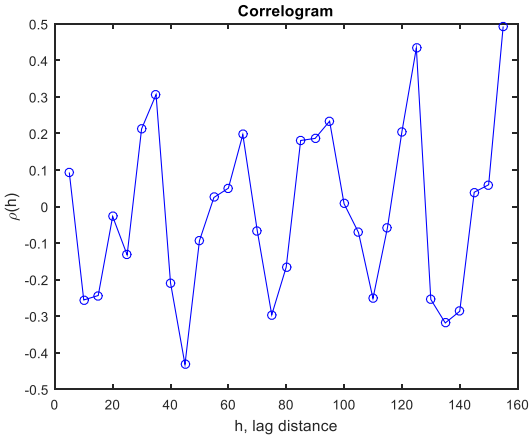


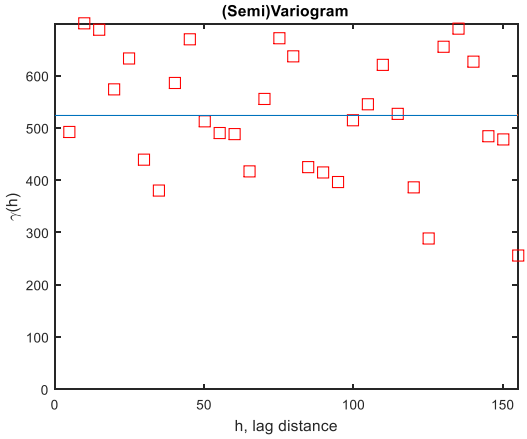


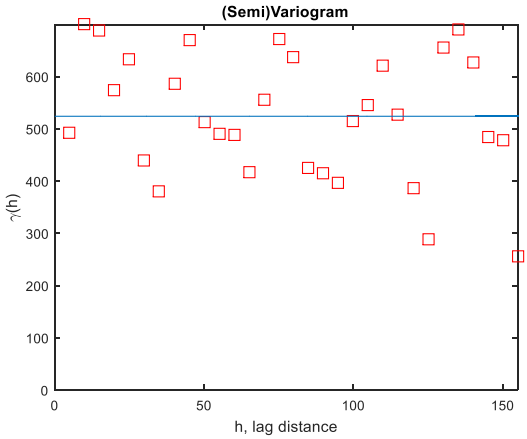
10m Shear wave Velocity with Trend Removed as 4th-Order Polynomial Function

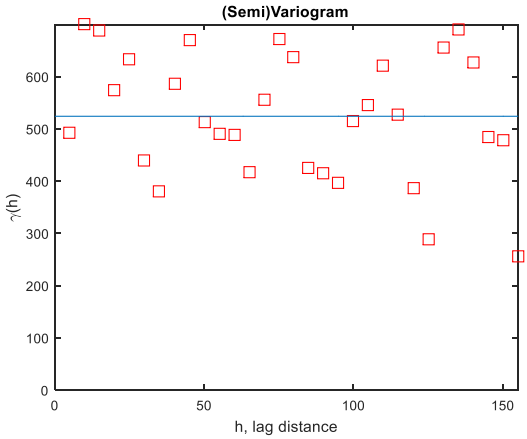




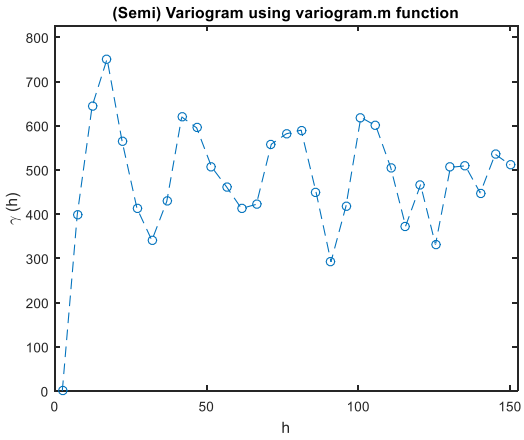


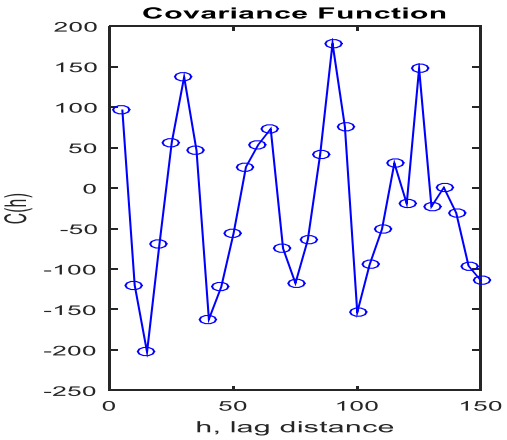


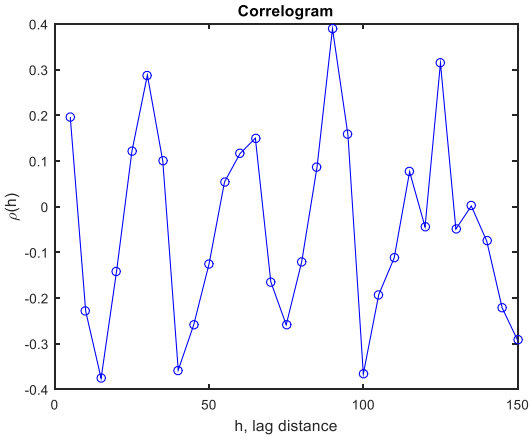


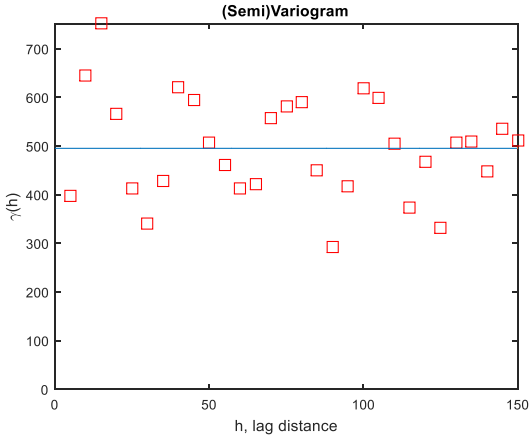


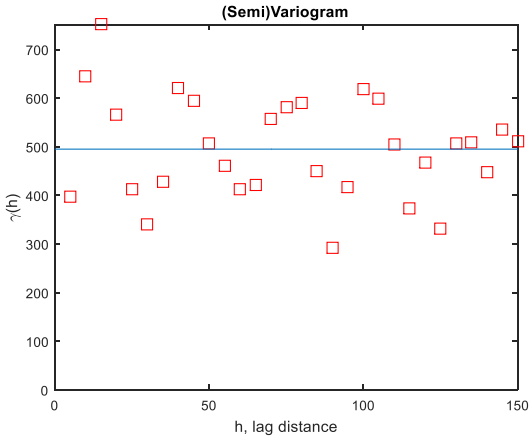
15m Shear wave Velocity with Trend Removed as 4th-Order Polynomial Function

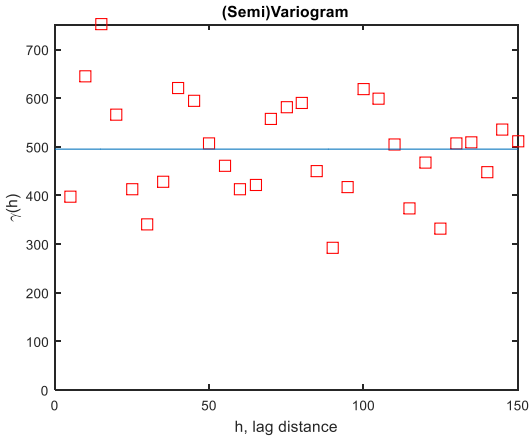




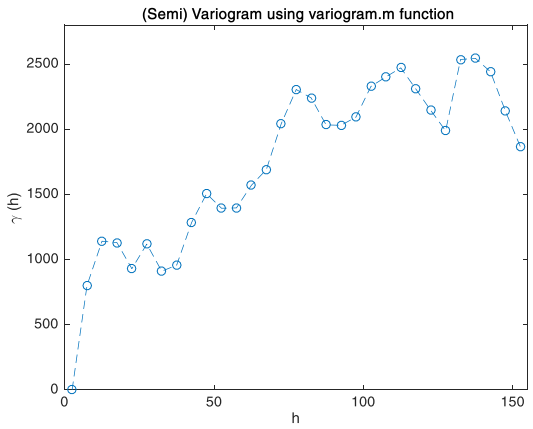


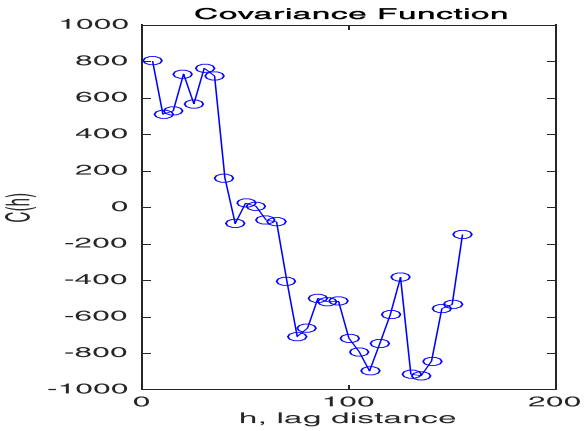


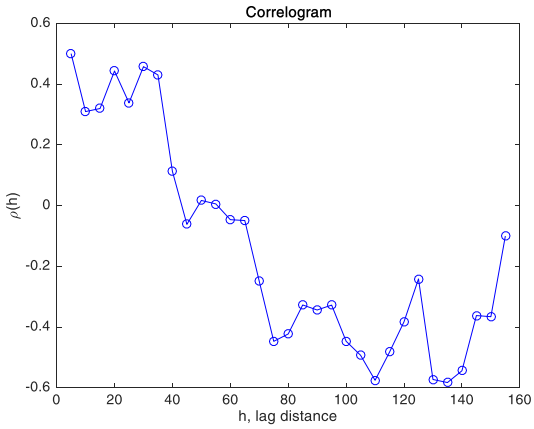


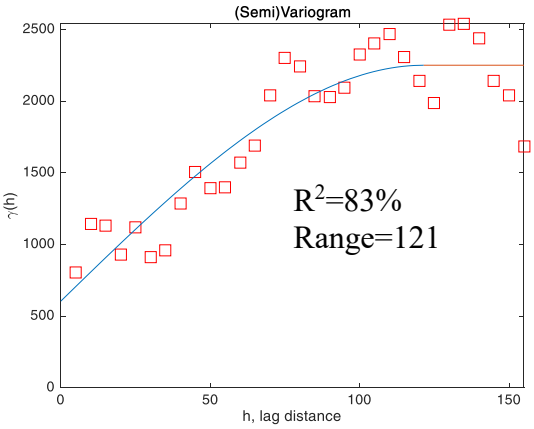


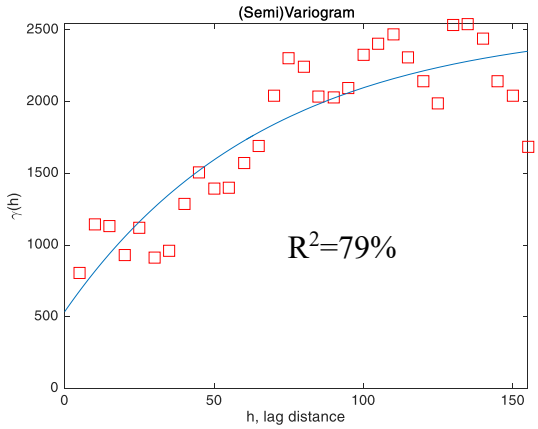
30m Shear wave Velocity based on Boore's 2004 study using 10m Shear wave Velocity

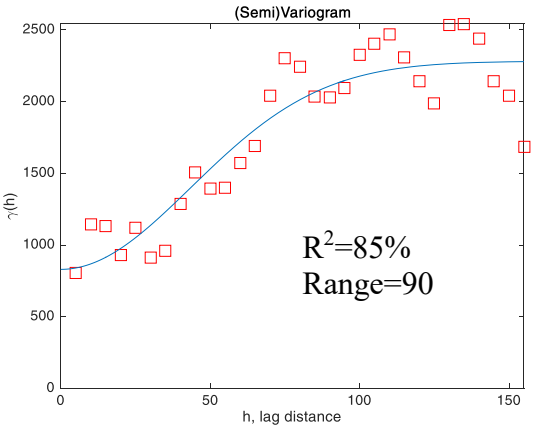












Appendix E: Personal MATLAB Code

This study used open-source MATLAB code (Schwanghart, 2010) to find and model semi-variograms, but MATLAB code to develop covariance functions and correlograms, verify semi-variogram calculations, and quantify spatial descriptors was written personally and can be referenced below.

```
clear

clc



%%DATA ENTRY%%
indicate_wave_type=1; %the type of wave that you are analyzing. Can
%choose either 1 for Rayleigh or 2 for Shear.
indicate_depth=10; %the depth at which you want to analyze. Can choose
%between 5m, 10m, or 15m (change this value as desired)

%The x-values are the coordinates along the north-south direction; where
%the point 0 is at the reference stake (see site map)
x=[-20;-15;-10;-5;0;5;10;15;20;25;30;35;40;45;50;55;60;65;70;75;80;85;90;
    95;100;105;110;115;120;125;130;135;140;145;150;155;160;165;170;175;
    180;185;190;195;220;225;230;235;240;245;250;255;260;265;270;275;280;
    285;290];
%The y-values are the coordinates along the east-west direction; notice
%that, for this project, data was collected in a straight line
y=[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;
    0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0];
%The z05R-values are the Rayleigh wave velocities at 5 meters depth; note
%that they are chosen from the HV curve.
%Note that values of "0" are not the actual phase velocities; this means
%that the location was not able to be measured. The code will delete these
%values automatically.
%Hill adjusted:
z05R=[180;174;179;182;213;245;260;0;250;174;295;261;264;258;255;263;176;
    284;0;222;260;0;213;163;164;247;243;210;180;169;0;225;218;195;248;
    252;230;189;164;200;225;151;243;152;274;277;257;279;271;241;265;293;
    317;279;278;267;229;273;273];
%Removed Trend with Linear Piecewise Function:
% z05R=[-50.182;-55.5715;-49.961;-46.3505;-14.74;17.8705;33.481;0;24.702;
%     -50.6875;70.923;37.5335;41.144;35.7545;33.365;41.9755;-44.414;64.1965;
%     0;3.4175;42.028;0;-3.751;-53.1405;-51.53;32.0805;28.691;-3.6985;
%     -33.088;-43.4775;0;13.7435;7.354;-15.0355;38.575;43.1855;21.796;
%     -18.5935;-42.983;-6.3725;19.238;-54.1515;38.459;-51.9305;0.72;3.97;
```

```
%     -15.78;6.47;-1.28;-31.03;-6.78;21.47;45.72;7.97;7.22;-3.53;-41.28;
%     2.97;3.22];
%Removed Trend with 4th-Order Polynomial Function:
% z05R=[39.5938991;11.65202707;-2.371285314;-15.68475721;1.509377006;
%     22.01536775;27.64395023;0;4.544255082;-75.53012822;42.82586886;
%     7.455405425;10.2081254;4.940157474;3.514115128;13.79909663;
%     -70.32931498;41.01094813;0;-13.34180618;28.76173598;0;-9.931306184;
%     -55.8924222;-51.02881467;35.59346253;34.91484031;4.882234338;
%     -22.5509549;-31.42484216;0;27.37725422;21.0054323;-1.902698761;
%     50.64517022;53.64783321;30.11056895;-12.954859;-40.53020234;
%     -7.59072795;13.89478206;-64.03596964;23.66120443;-71.9630235;
%     24.32227059;22.19689588;-2.713601115;14.69333015;2.526724979;
%     -31.09789652;-10.05852947;15.7733158;38.53261374;0.360823598;
%     0.405889396;-8.177760058;-42.22921117;7.418934444;14.94055995];
%10m Point values (not averaged):
% z10R=[325;340;365;320;345;340;345;375;350;365;385;405;380;375;335;345;305;
%     345;350;365;345;300;325;320;330;335;335;300;285;315;330;370;360;280;
%     305;310;300;300;300;330;280;290;320;320;360;365;360;385;360;365;340;
%     480;400;395;390;350;365;375;425];

%These are the old 10 m phase velocity values:
% z10old=[325;350;375;355;375;340;360;380;370;355;385;395;405;410;370;370;325;370;
%     365;0;380;280;325;345;345;355;350;310;300;325;310;325;360;275;325;330;
%     315;320;305;285;275;300;340;320;365;370;350;375;380;385;320;395;415;0;
%     365;335;365;365;310];

%The z05S-values are the Shear wave velocities at 5 meters depth; they are
%pulled from the shear wave velocity model given after inversion.
%Note that values of "0" are not the actual phase velocities; this means
%that the location was not able to be measured. The code will delete these
%values automatically.
%Hill adjusted:
z05S=[193;211;236;234;255;259;259;0;248;240;293;295;298;293;266;262;190;
    324;0;254;268;0;238;209;214;269;255;230;188;209;0;245;249;239;
    240;232;220;192;193;206;234;178;232;189;301;313;287;299;325;301;286;
    359;361;308;298;301;286;293;309];
%Non-hill adjusted:
% z05S=[193;211;236;234;255;259;259;0;248;240;293;295;298;293;266;262;190;
%     324;0;254;268;0;238;209;214;269;255;230;188;209;0;245;249;239;
%     240;232;220;192;193;206;234;178;232;189;255;270;255;307;282;257;280;
%     299;332;265;250;264;244;261;259];
%Removed Trend with Linear Piecewise Function:
% z05S=[-69.178;-50.0885;-23.999;-24.9095;-2.82;2.2695;3.359;0;-5.462;-12.3725
% 41.717;44.8065;48.896;44.9855;19.075;16.1645;-54.746;80.3435;0;12.5225;
% 27.612;0;-0.209;-28.1195;-22.03;34.0595;21.149;-2.7615;-43.672;-21.5825;0;
% 16.5965;21.686;12.7755;14.865;7.9545;-2.956;-29.8665;-27.777;-13.6875;
% 15.402;-39.5085;15.581;-26.3295;-54.782;-39.5925;-54.403;-2.2135;-27.024;
% -51.8345;-28.645;-9.4555;23.734;-43.0765;-57.887;-43.6975;-63.508;-46.3185;
% -48.129];
%Removed Trend with 4th-Order Polynomial Function:
```

```
% z05S=[1.60847366;1.76722279;11.43151082;-3.58025555;6.55601838;1.67011531;
%     -5.40249406;0;-25.7375089;-36.30053297;15.34849946;17.07349909;
%     20.74406462;17.23548275;-7.57127182;-8.78953639;-77.52696026;
%     60.11449527;0;-1.84536067;16.37784526;0;-5.16552158;-30.07344895;
%     -21.15696202;37.52748191;26.92911354;5.00285157;-34.2906973;
%     -10.98523837;0;28.31632169;33.27345222;23.75164935;24.74564778;
%     16.25587021;3.28842734;-26.14488213;-27.0265715;-16.33346607;
%     8.96329686;-50.10173201;0.51168602;-46.15052235;-0.5903657;
%     10.37665623;-8.45119734;40.01750429;11.87987982;-15.76126405;
%     5.20256738;22.88555681;55.40757494;-11.10581953;-24.5233799;
%     -7.70817147;-23.51757154;-0.80326941;4.58873362];

%The z12R-values are the Rayleigh wave velocities at 12 meters depth; note
%that they are just "eye-balled" values from the HV curve.
%Note that values of "0" are not the actual phase velocities; this means
%that the location was not able to be measured. The code will delete these
%values automatically.
%Non-Hill adjusted:
z10R=[210;210;226;224;249;263;280;332;275;222;309;297;293;294;282;280;208;
    313;305;262;284;290;243;202;209;265;258;231;210;210;303;255;251;225;
    256;253;242;213;197;230;243;182;253;188;294;299;286;310;302;275;283;
    332;341;305;302;289;265;298;299];
%Removed Trend with Linear Piecewise Function
% z10R=[-61.77;-60.8;-43.83;-44.86;-18.89;-3.92;14.05;67.02;10.99;-41.04;
%     46.93;35.9;32.87;34.84;23.81;22.78;-48.25;57.72;50.69;8.66;31.63;38.6;
%     -7.43;-47.46;-39.49;17.48;11.45;-14.58;-34.61;-33.64;60.33;13.3;10.27;
%     -14.76;17.21;15.18;5.15;-22.88;-37.91;-3.94;10.03;-50;21.97;-42.06;
%     -5.72;-0.57;-13.42;10.73;2.88;-23.97;-15.82;33.33;42.48;6.63;3.78;
%     -9.07;-32.92;0.23;1.38];
%Removed Trend with 4th-Order Polynomial Function:
% z10R=[16.16226914;-2.891364489;-3.340443671;-19.37100678;-6.163364854;
%     -1.892101543;7.273926843;53.17159136;-8.356509593;-64.46205023;
%     20.70902256;8.016489276;4.32585776;6.508363213;-3.559031801;
%     -2.993637363;-71.90703619;36.59491634;32.41209224;-6.549909138;
%     19.62023891;29.83959046;-12.96907306;-49.87724286;-38.95068279;
%     20.75057064;17.17220831;-6.734351567;-25.01196341;-22.69775428;
%     72.17687613;25.58625546;22.51043873;-3.064791684;27.85207395;
%     24.25827271;12.15676905;-18.44374521;-36.52885093;-6.078401567;
%     3.933476753;-60.46161473;6.773652586;-62.31766534;20.2348399;
%     20.40891767;2.718167925;22.2514655;10.10341257;-20.62566132;
%     -15.82969926;30.60308299;37.79019706;0.85488192;-2.073896093;
%     -13.86144329;-35.36733861;1.554566346;8.056147348];
% z10R=flipud(z10Rcorrect) %to check the reverse direction
%12m Point values (not averaged):
% z12R=[365;375;400;345;370;380;375;410;360;380;420;435;410;395;360;380;340;
%     365;375;395;365;340;340;340;375;365;365;330;320;350;360;0;395;305;
%     330;340;330;335;335;365;340;310;360;340;360;365;360;385;360;365;340;
%     480;400;395;390;350;365;375;425];
%These are the old 12m phase velocity values:
% z12old=[365;395;400;360;395;365;375;400;390;380;420;420;420;420;380;390;355;380;
```

```
%     390;375;415;305;345;360;380;380;380;330;335;365;355;0;390;315;355;355;355;
%     340;345;325;310;330;375;360;380;380;380;400;405;395;335;400;420;420;410;
%     355;385;385;335];

%The z10S-values are the Shear wave velocities at 10 meters depth; they are
%pulled from the shear wave velocity model given after inversion.
%Note that values of "0" are not the actual phase velocities; this means
%that the location was not able to be measured. The code will delete these
%values automatically.
%Hill adjusted:
z10S=[302;334;337;350;312;310;334;376;345;329;368;385;346;375;363;343;283;
    368;353;350;355;328;298;289;326;332;326;311;294;281;319;328;346;292;
    293;296;307;288;297;325;289;274;331;297;371;375;375;401;386;368;334;
    454;397;399;372;353;353;378;408];
%Non-hill adjusted:
% z10S=[302;334;337;350;312;310;334;376;345;329;368;385;346;375;363;343;283;
%     368;353;350;355;328;298;289;326;332;326;311;294;281;319;328;346;292;
%     293;296;307;288;297;325;289;274;331;297;348;355;345;377;367;346;311;
%     423;386;369;348;337;328;350;369];
%Removed Trend with Linear Piecewise Function:
% z10S=[-48.374;-15.178;-10.982;3.214;-33.59;-34.394;-9.198;33.998;4.194;
%     -10.61;29.586;47.782;9.978;40.174;29.37;10.566;-48.238;37.958;24.154;
%     22.35;28.546;2.742;-26.062;-33.866;4.33;11.526;6.722;-7.082;-22.886;
%     -34.69;4.506;14.702;33.898;-18.906;-16.71;-12.514;-0.318;-18.122;
%     -7.926;21.27;-13.534;-27.338;30.858;-1.946;-32.148;-25.355;-35.562;
%     -3.769;-13.976;-35.183;-70.39;41.403;4.196;-13.011;-34.218;-45.425;
%     -54.632;-32.839;-14.046];
%Removed Trend with 4th-Order Polynomial Function:
% z10S=[0.452582866;20.01986364;12.46899293;16.6599005;-28.54307087;
%     -36.2711654;-16.65121432;22.19436414;-10.85243524;-27.90520468;
%     10.92687659;28.53904234;-9.16906065;21.7066274;12.07457925;
%     -5.152319304;-62.0567695;26.28294044;14.79293528;15.40375281;
%     24.05034381;0.672072033;-25.78728573;-31.37953971;8.847912874;
%     17.85208779;14.59441382;2.04073273;-12.8387007;-24.06921869;
%     15.32825953;25.33722773;43.94559269;-9.854325813;-9.065795006;
%     -6.687669114;3.285610638;-17.13598397;-9.938068178;15.8981558;
%     -22.60410125;-40.41721558;13.49084961;-23.84345592;8.434107842;
%     11.6581513;-1.981463534;26.58701712;13.43976003;-10.34265502;
%     -47.67523527;62.53142508;24.37114478;6.942155615;-13.65289764;
%     -23.30695721;-29.90855231;-4.341799187;19.51359895];

%The z15R-values are the Rayleigh wave velocities at 15 meters depth; note
%that they are just "eye-balled" values from the HV curve.
%Note that values of "0" are not the actual phase velocities; this means
%that the location was not able to be measured. The code will delete these
%values automatically.
%Non-Hill adjusted:
z15R=[245;250;268;254;278;289;306;356;306;258;338;329;324;319;312;307;239;
    337;326;297;308;0;268;236;245;292;286;257;242;246;321;0;0;0;277;278;
    269;244;230;0;0;214;281;221;318;321;315;335;328;303;302;0;360;335;
```

```
    329;310;295;325;0];
%Removed Trend with Linear Piecewise Function:
% z15R=[-57.476;-51.3545;-32.233;-45.1115;-19.99;-7.8685;10.253;61.3745;
%     12.496;-34.3825;46.739;38.8605;34.982;31.1035;25.225;21.3465;-45.532;
%     53.5895;43.711;15.8325;27.954;0;-9.803;-40.6815;-30.56;17.5615;12.683;
%     -15.1955;-29.074;-23.9525;52.169;0;0;0;12.655;14.7765;6.898;-16.9805;
%     -29.859;0;0;-42.4945;25.627;-33.2515;-4.03;-0.905;-6.78;13.345;6.47;
%     -18.405;-19.28;0;38.97;14.095;8.22;-10.655;-25.53;4.595;0];
%Removed Trend with 4th-Order Polynomial Function:
% z15R=[14.22515578;1.354720969;4.011223226;-22.98865739;-9.82253464;
%     -7.662316129;2.325796718;46.98130865;-6.850569429;-57.31892042;
%     21.43287893;12.26715804;8.051952528;4.661004154;-0.026239139;
%     -2.124623237;-66.74328786;35.01433346;28.05051333;3.273230568;
%     18.59617013;0;-12.77401304;-40.61123498;-27.63643301;23.0926087;
%     20.52381213;-5.389194533;-17.68707697;-11.40479466;65.42839907;0;0;0;
%     23.88493777;24.23360479;14.07749289;-12.57270804;-28.70060195;0;0;
%     -53.70088856;9.538529485;-54.53217265;19.43438762;17.76597497;
%     7.261703262;23.01214876;12.11359391;-16.33197268;-20.21685621;0;
%     34.15473588;8.655131925;3.204051876;-14.06227863;-26.00192778;
%     8.532742417;0];
%15m Point values (not averaged):
% z15R=[425;415;445;365;410;415;410;435;400;405;465;480;445;425;395;415;385;
%     390;415;440;390;0;365;380;420;405;410;365;365;400;410;0;0;0;
%     360;380;390;380;395;0;0;360;415;370;395;395;415;410;385;400;375;
%     0;430;435;430;375;410;420;0];
%These are the old 15m phase velocity values:
% z15old=[455;435;440;365;430;405;410;430;415;405;440;440;445;440;410;420;400;
%     400;390;415;445;0;360;380;425;420;420;370;380;415;0;0;0;0;
%     380;395;400;375;380;380;0;365;425;385;405;405;415;440;415;410;355;
%     400;435;465;455;385;0;435;0];

%The z15S-values are the Shear wave velocities at 15 meters depth; they are
%pulled from the shear wave velocity model given after inversion.
%Note that values of "0" are not the actual phase velocities; this means
%that the location was not able to be measured. The code will delete these
%values automatically.
%Hill adjusted:
z15S=[384;422;429;360;367;383;377;424;395;378;418;429;415;415;412;399;354;
    418;407;417;391;0;346;364;404;404;392;347;376;362;388;0;0;0;
    345;373;369;345;362;0;0;334;386;367;410;408;429;442;418;404;374;
    0;422;455;422;379;404;429;0];
%Non-hill adjusted:
% z15S=[384;422;429;360;367;383;377;424;395;378;418;429;415;415;412;399;354;
%     418;407;417;391;0;346;364;404;404;392;347;376;362;388;0;0;0;
%     345;373;369;345;362;0;0;334;386;367;396;397;410;429;407;392;360;0;
%     408;435;400;370;386;412;0];
%Removed Trend with Linear Piecewise Function:
% z15S=[-23.526;15.578;23.682;-44.214;-36.11;-19.006;-23.902;24.202;-3.694;
%     -19.59;21.514;33.618;20.722;21.826;19.93;8.034;-35.862;29.242;19.346;
%     30.45;5.554;0;-37.238;-18.134;22.97;24.074;13.178;-30.718;-0.614;
```

```
%     -13.51;13.594;0;0;0;-24.99;4.114;1.218;-21.678;-3.574;0;0;-28.262;
%     24.842;6.946;-20.588;-19.3525;-6.117;13.1185;-8.646;-23.4105;-55.175;0;
%     -6.704;20.5315;-14.233;-43.9975;-27.762;-1.5265;0];
%Removed Trend with 4th-Order Polynomial Function:
% z15S=[4.134949991;35.50529324;36.95500408;-36.59690443;-33.22887974;
%     -20.01682981;-28.03412309;17.64841148;-12.03751556;-29.15765415;
%     11.22478526;23.0491317;10.25725375;11.79355944;10.60499632;
%     -0.358948549;-43.14624863;23.19766214;14.62988931;27.11007792;
%     3.600412521;0;-36.52704459;-16.2077697;25.99578439;28.05849972;
%     17.95779784;-25.32636019;5.188526173;-7.512503018;19.55813179;0;0;0;
%     -21.60221555;6.229314981;1.811316305;-22.85085553;-6.749304962;0;0;
%     -38.7452709;11.54091474;-9.331606967;8.823695903;7.573596112;
%     18.38004632;35.28395556;11.32877241;-5.439515102;-38.97237942;0;
%     6.874969235;33.3649348;-1.690171232;-31.22912382;-14.18815842;
%     13.49902903;0];

%The z20S-values are the Shear wave velocities at 20 meters depth; they are
%pulled from the shear wave velocity model given after inversion.
%Note that values of "0" are not the actual phase velocities; this means
%that the location was not able to be measured. The code will delete these
%values automatically.
%Hill adjusted:
z20S=[0;0;0;364;434;443;429;454;424;410;449;452;0;438;436;434;0;
    449;0;0;411;0;0;0;459;453;436;0;0;0;0;0;0;0;
    378;429;410;0;0;0;0;0;0;0;0;438;0;469;0;0;0;
    0;0;0;0;0;0;0;0];
%Non-hill adjusted:
% z20S=[0;0;0;364;434;443;429;454;424;410;449;452;0;438;436;434;0;
%     449;0;0;411;0;0;0;459;453;436;0;0;0;0;0;0;0;
%     378;429;410;0;0;0;0;0;0;0;426;427;0;460;0;424;0;0;451;0;454;
%     0;0;0;0];

%Choosing which z-values to use
if indicate_wave_type==1
    if indicate_depth==5
    z=z05R;
    else
        if indicate_depth==10
            z=z10R;
        else
            if indicate_depth==15
                z=z15R;
            else
        fprintf('Error: the depth needs to be at 5, 10, or 15 meters.\n')
            end
        end
    end
else
    if indicate_wave_type==2
        if indicate_depth==5
```

```
        z=z05S;
        else
            if indicate_depth==10
                z=z10S;
            else
                if indicate_depth==15
                    z=z15S;
                else
                    if indicate_depth==20
                        z=z20S;
                    else
                        fprintf('Error: the depth needs to be at 5, 10, 15, or 20
meters.\n')
                    end
                end
            end
        end
    else
        fprintf('Error:Wave type needs to be either "1" for Rayleigh or "2" for Shear.
\n')
    end
end

%%(SEMI) VARIOGRAM CALCULATION%%

%Note that this analysis will (most likely) only work for one-dimensional
%arrays since that is the situation I am scripting for


t1=numel(z); %t gives the original number of elements in z
h=5; %h is the initial lag distance in meters
h_increment=5; %increment to the next lag (assumes constant increment)
gamma_h=[]; %initial (semi)variogram matrix
H=[]; %initial lag vector
C_h=[]; %intial covariance-function vector
rho_h=[]; %initial correlation-function vector
Data_pairs_for_each_lag=[]; %initial data-pair vector

%%Making sure that the location vector, x, has ascending values:

if x(1:(t1-1)) > x(2:t1)
    fprintf('Error. Values in the location matrix must be ascending; results will be
erroneous elsewise.')
end

%%Deleting any z=0 values (and corresponding x and y values):

i0=1; %these 3 values are just initial values for the loop's use
itemsdeleted=0;
rowswithzeros=[];
```

```
elements_in_z=numel(z);
while i0 <= elements_in_z
if z(i0)==0
    z(i0)=[];
    x(i0)=[];
    y(i0)=[];
    itemsdeleted=itemsdeleted+1;
    rowswithzeros=[rowswithzeros;i0];
    elements_in_z=numel(z);
else
    i0=i0+1;
    continue
end
end

%%Deleting any z values that are 2 standard deviations away from the mean
%%(and corresponding x and y values):
%%Note that this didn't seem to help this study's variogram/covariance fit,
%%so it is commented out.

% count=1; %value is just an initial value for the loop's use
% elements_in_z=numel(z);
% while count <= elements_in_z
% if z(count)>mean(z)+2*std(z)
%     z(count)=[];
%     x(count)=[];
%     y(count)=[];
%     itemsdeleted=itemsdeleted+1;
%     elements_in_z=numel(z);
% else
%     if z(count)<mean(z)-2*std(z)
%         z(count)=[];
%         x(count)=[];
%         y(count)=[];
%         itemsdeleted=itemsdeleted+1;
%         elements_in_z=numel(z);
%     else
%         count=count+1;
%         continue
%     end
% end
% end

if itemsdeleted > 0
    fprintf('Be aware that %d element(s) in z that equal(s) zero have been
deleted.\n',itemsdeleted)
    fprintf('-This is for a lag distance of %d meters.\n',h)
end

%%Making the location vector, x, have all positive values that start at 0:
```

```
if x(1)<0
    xpos=x+abs(x(1));
else
    xpos=x-x(1);
end

%%Where to bracket the lag distance (typically half of the total distance):
maxlag=0.5*(max(x)-min(x));

while h <= maxlag %looping until data for all lag values are found

%%Finding the number of data pairs, N, for the specified lag and seeing
%%which rows correspond to the specified lag:

t2=numel(z); %t2 is the new number of elements in z, since some may have been deleted
i1=1; %the rest of the values are initial values for the loop's use
rowsbeingused=[];
xabsent=0;
N=0;
xpos1=xpos;
xpos2=xpos;
xlag=xpos;

while numel(xlag) > 0
    xpos1=[zeros(i1,1);xpos]; %shifts the vector down by one row
    xpos2=[xpos;zeros(i1,1)]; %shifts the vector up by one row
    xlag_too_big=abs(xpos1-xpos2); %finds the difference (lag value)between
    %the shifted vectors
    a1=i1+1;
    a2=numel(xpos);
    xlag=xlag_too_big(a1:a2); %since we added rows, this crops the vector back to the
size we want
    a3=find(xlag==h);%searches for the rows that gave our designated lag distance
    N=N+numel(a3); %if there was one, then this adds it to the amount of data pairs
    if numel(a3)>0;
        rowsbeingused=[rowsbeingused;a3]; %tells which row had the correct lag
distance
    end
    i1=i1+1;
end
rowsbeingused=sort(rowsbeingused); %puts the rows being used in ascending order

Data_pairs_for_each_lag=[Data_pairs_for_each_lag;h,N]; %gives the amount of
%data pairs listed next to the corresponding lag value

%%Finding which rows do not correspond to the specified lag:

totalrows=[1:t2]'; %total rows in the location vector, x
difference_in_dimensions=size(totalrows)-size(rowsbeingused);%amount of rows needed
```

```
%to make the dimensions the same as the vector with the total number of
%rows in "x"
rows_to_add=zeros(difference_in_dimensions(1),1); %creates the rows to add
rowsbeingused_samedimension=[rowsbeingused;rows_to_add]; %adds the rows
rowsbeingskipped=[]; %these next 3 values are intial values for the loop's use
i2=1;
i3=1;
while i3<=t2
    if rowsbeingused_samedimension(i2)==i3
        i2=i2+1;
        i3=i3+1;
        continue
    else
        rowsbeingskipped=[rowsbeingskipped;i3];
    end
    i3=i3+1;
end
numbermissing=numel(rowsbeingskipped(:,1)); %the number of rows that didn't satisfy
the specified lag distance

%%Creating a vector of the squared difference of each lag value:
RBS=sort(rowsbeingskipped); %rows being skipped in ascending order
i4=1; %these values are initial values for the loop's use
i5=1;
V=0;
C1=0;
m_minus_h=0;
m_plus_h=0;
v_m_h_2=0;
v_p_h_2=0;
H_scatterplot_x=[];
H_scatterplot_y=[];
while i5 <= t2
    if RBS(i4)==i5
        i4=i4+1;
        i5=i5+1;
        continue
    else
        xposref=xpos(i5)+h; %referencing the location of the other data
        %point that corresponds with the specified lag distance
        corresp_row=find(xpos==xposref); %finds the corresponding row
        V=V+(z(corresp_row)-z(i5))^2; %sums the squared differences of each
        %of the z values that correspond to the specified lag distance
        C1=C1+z(corresp_row)*z(i5); %sums the product of the z values that
        %correspond to the specified lag distance
        m_minus_h=m_minus_h+(z(corresp_row)/N); %average value for first location
value
        m_plus_h=m_plus_h+(z(i5)/N); %average value for second location values
        v_m_h_2=v_m_h_2+(z(corresp_row))^2; %squared sum  for the first location
        v_p_h_2=v_p_h_2+(z(i5))^2; %squared sum for the second location
```

```
    end
    i5=i5+1;
end

%%Creating the semivariogram point corresponding to the specified lag by
%%dividing by the number of data pairs and halving the value:

H=[H;h]; %creating lag vector
gamma_h=[gamma_h;V/(2*N)]; %creating (semi)variogram vector

%%Creating the covariance function point

C_h=[C_h;(C1-N*(m_minus_h*m_plus_h))/N]; %creating the covariance function vector
%Notice that the multiplication of the mean values is multiplied by the
%number of pairs. This is because they are a constant value subtracted from
%the equation N amount of times.

%%Creating the correlation function point

sigma_minus_h=sqrt((v_m_h_2/N)-(m_minus_h)^2); %standard deviation for the first
location
sigma_plus_h=sqrt((v_p_h_2/N)-(m_plus_h)^2); %standard deviation for the second
location
rho_h=[rho_h;((C1-N*(m_minus_h*m_plus_h))/N)/(sigma_minus_h*sigma_plus_h)]; %creating
%the correlation function vector

h=h+h_increment; %proceeding to the next lag value
end

%Amount of data pairs
display(Data_pairs_for_each_lag)

%semivariogram values:
semivariogram_data=table(H,gamma_h)

figure %create new figure
subplot(1,2,1)
plot(H,gamma_h,'ko-')
title('(Semi)Variogram')
xlabel('h, lag distance')
ylabel('\gamma(h)')

%covariance function:
covariance_function_data=table(H,C_h)

subplot(1,2,2)
plot(H,C_h,'bo-')
title('Covariance Function')
xlabel('h, lag distance')
ylabel('C(h)')
```

```
%Correlogram:
correlation_function_data=table(H,rho_h)

figure
plot(H,rho_h,'bo-')
title('Correlogram')
xlabel('h, lag distance')
ylabel('\rho(h)')

%can plot both the semivariogram and covariance function on the same plot
%by uncommenting the following:
% figure %creates new figure
% [hAx,hLine1,hLine2] = plotyy(H,gamma_h,H,C_h);
% title('(Semi)Variogram and Covariance Function')
% xlabel('h, lag distance')
% ylabel(hAx(1),'\gamma(h)')
% ylabel(hAx(2),'C(h)')
% set(hLine1,'LineStyle','-')
% set(hLine2,'LineStyle','o')

%Fitting a function to the (semi)variogram plot using the open source
%variogramfit function
figure
[range,sill,nugget,Rs]=variogramfit(H,gamma_h,155,1700,[],'model','spherical','nugget'
,500)
%sets the largest lag value, the initial sill value, the model type
%(spherical, exponential, and Gaussian were used in this study), and the
%initial nugget effect value
title('(Semi)Variogram')
xlabel('h, lag distance')
ylabel('\gamma(h)')

%Using the open-source variogram function
    figure
%     subplot(2,1,1)
%     hist(z,10)
%     ylabel('frequency'); xlabel('z')
%     title('histogram of z-values')
%     subplot(2,1,2)
    d = variogram([x y],z,'plotit',true,'nrbins',31);
    title('(Semi) Variogram using variogram.m function')
```